

Rajapidennyksen produktiivisuuden foneettinen tarkastelu e-loppuisissa nomineissa

Nina Rainio
Pro gradu -tutkielma
Fonetiikka
Humanistinen tiedekunta
Helsingin yliopisto
Marraskuu 2020

Tiedekunta/Osasto – Fakultet/Sektion – Faculty Humanistinen tiedekunta		Laitos – Institution – Department Digitaalisten ihmistieteiden osasto	
Tekijä – Författare – Author Nina Rainio			
Työn nimi – Arbetets titel – Title Rajapidennyksen produktiivisuuden foneettinen tarkastelu e-loppuisissa nomineissa			
Oppiaine – Läroämne – Subject Fonetiikka			
Työn laji / Ohjaajat – Arbetets art / Handledare – Level / Instructor Pro gradu -tutkielma / Martti Vainio Katri Hiovain		Aika – Datum – Month and year Marraskuu 2020	Sivumäärä – Sidoantal – Number of pages 57
Tiivistelmä – Referat – Abstract <p>Äänteen pitenemistä sanan rajalla yhtäjaksoisessa puheessa nimitetään rajapidennykseksi. Tämän ortografiassa näkymättömän, puheessa kuuluvan ilmiön aiheuttavat tietyt, kielenmuutoksen seurauksena nykyään vokaaliloppuiset sananmuototyytit. Tässä työssä tarkastellaan rajapidennyksen foneettista toteutumista ilmiölle tunnusomaisessa e-loppuisten nominien muototyytissä. Tavoitteena on selvittää, onko ilmiö produktiivinen, ja tätä tarkastellaan e-loppuisia <i>vene</i>-tyypin sanoja mukailevien pseudosanojen avulla.</p> <p>Tutkimusta varten toteutettiin puheentuottokoe, jossa 14 äidinkieliseltä suomen kielen puhujalta äänitettiin 2412 koelauseita ilmiön mahdollisesti laukaisevine e-loppuisine herätesanoineen. Herätesanat olivat neljäntyyppisiä: oikeita e-loppuisia <i>vene</i>-tyypin sanoja, joiden oletettiin laukaisevan ilmiö, nimiä, joiden yhteydessä ilmiön toteutumista ei odotettu, sekä tutkimusta varten kehitettyjä epäsanoina ja epälainasanoja. Herätesanat olivat konsonantti-vokaali-rakenteeltaan kahdenlaisia, ja sanatyypit erosivat toisistaan ensimmäiseltä äänteeltään. Jokainen herätesana luettiin kolmessa eri kehyslauseessa, kerran ennen kutakin kohdeäännettä, jotka olivat [h], [m] ja [p]. Herätesanan jälkeisen kohdeäänteen toteutunut kesto segmentoitui ja mitattiin. Sanatyypin, sanan rakenteen ja kohdeäänteen vaikutusta rajapidennyksen toteutumiseen analysoitiin sekä parametrisin että epäparametrisin tilastollisin menetelmin. Kohdeäänteille tehtiin myös niiden havaittuun keston perustuva auditiivinen luokittelu.</p> <p>Herätesanoissa ilmeni odotetusti eroja ilmiön laukaisemisessa. Tulokset osoittivat, että sanatyypillä oli tilastollisesti erittäin merkitsevä vaikutus ilmiön toteutumiseen, ja sanan rakenteen vaikutus osoittautui tilastollisesti merkitseväksi tekijäksi normalisoiduissa epälainasanoissa. Ennako-oletusten mukaisesti rajapidennys toteutui oikeiden sanojen yhteydessä kaikkein useimmin ja nimien yhteydessä harvimminkin. Epäsanat osoittautuivat läheisesti oikeiden sanojen kaltaisiksi, epälainasanat muistuttivat puolestaan nimiä. Kohdeäänteistä lähtökohtaisesti pitkäkestoisin [p] osoittautui myös altteimmaksi pitenemään, ja [h]:n rajapidennystendenssi oli aiempien tutkimusten mukaisesti vähäisempi kuin muilla äänteillä.</p> <p>Rajapidennys osoittautui tämän tutkimuksen perusteella produktiiviseksi ilmiöksi e-loppuisissa nomineissa, sillä tutkimusta varten keksityt epäsanat laukaisivat sen lähes yhtä lailla kuin oikeat <i>vene</i>-tyypin sanat. Tämä puoltaa näkemystä siitä, että ilmiö on elinvoimainen, ja sen paikoittainen toteutumattomuus ei indikoi sen katoamista suomesta. Tutkimus vahvisti kuitenkin aikaisempia huomioita siitä, että puhuja- ja tilannekohtaista vaihtelua on runsaasti.</p>			
Avainsanat – Nyckelord – Keywords fonetiikka, rajapidennys, rajageminaatio, loppukahdennus, epäsanat			
Säilytyspaikka – Förvaringställe – Where deposited Helsingin yliopiston kirjasto – Helda / E-thesis			
Muita tietoja – Övriga uppgifter – Additional information			

Sisällysluettelo

1	Johdanto	1
1.1	Rajapidennys	2
1.1.1	Määrittelyä	3
1.1.2	Merkityseron aiheuttava sandhi	3
1.2	Ilmiön nimityksestä	4
1.3	Rajapidennyksen aiheutuminen	7
1.3.1	Rajapidennyksen aiheuttavat muotoryhmät	7
1.3.2	Kielihistoriallista taustaa	8
1.4	Rajapidennyksen toteutuminen ja aluekohtaisuus	9
1.5	E-loppuiset nominit: [hernek:eit:o] ja [nuk:ekoti]	10
1.6	Suomen kvantiteetti	12
1.7	Paino	13
1.8	Äänteiden kesto ja laatu	14
1.8.1	Äänteiden laatu ja rajapidennys	15
1.8.2	Tapaus [h]	16
2	Tutkimusmenetelmät ja aineisto	17
2.1	Koehenkilöt	17
2.2	Koemateriaali	18
2.2.1	Herätesanat	18
2.2.2	Kohdeäänteet	20
2.2.3	Lauseet	20
2.2.4	Pilottikoe, ensimmäiset äänitykset ja lukumateriaalimuutos	23
2.3	Tutkimuskysymykset	23
2.4	Äänitys	25
2.5	Aineiston käsittely	26
2.5.1	Akustinen analyysi	26
2.5.2	Auditiivinen analyysi	28
2.6	Tilastolliset testit	29
3	Tulokset	32
3.1	Sanatyypin vaikutus rajapidennyksen toteutumiseen	32
3.1.1	Lineaarinen regressioanalyysi: z-muunnettu aineisto	32

3.1.2	Lineaarinen regressioanalyysi: absoluuttiset kestot	34
3.1.3	Kruskall-Wallis testi: z-muunnettu aineisto.....	35
3.1.4	Kruskall-Wallis testi: absoluuttiset kestot.....	36
3.2	Sanan rakenteen vaikutus rajapidennyksen toteutumiseen.....	36
3.2.1	Varianssianalyysi: z-muunnettu aineisto.....	37
3.2.2	Kruskall-Wallis testi: z-muunnettu aineisto.....	38
3.3	Kohdeäänteet	38
3.4	Auditiivinen luokittelu	40
4	Pohdinta	44
4.1	Hypoteesien toteutumisen tarkastelu	45
4.2	Ennakko-oletusten toteutumisen tarkastelu	46
4.3	Koehenkilöt	48
4.4	Muita huomioita tutkimuksesta	49
4.5	Tutkimuksen arviointia.....	50
4.6	Mietteitä jatkotutkimuksesta	52
5	Yhteenveto	53
	Lähteet.....	54

1 Johdanto

Kansanomaisen käsityksen mukaan suomen lausuminen on helppoa, sillä suomea puhutaan, niin kuin sitä kirjoitetaan. Puheen ja kirjoituksen suhde suomessa on toki melko säännöllinen, ja moniin kieliin nähden ortografia mukailee paljolti puhuttua kieltä, mutta säännöllisyydessä esiintyy kuitenkin poikkeamia. Suomi on produktiivinen kieli yhdyssanojen ja johdosten muodostamisessa (Koivisto, 2013, 232–234) sekä sanojen muokkaamisessa kirjoitusasuun sujuvammin puheeseen sopiviksi ja korvaamisessa toisella sanalla. Kirjoitetun ja puhutun kielen ero tulee ilmi myös vaihto-oppilaiden haastatteluista; jopa hyvin suomen hallitsevilla on sanastosta ja sanojen lyhentämisestä johtuen haasteita ymmärtää puhuttua kieltä (Lauranto & Vehkanen, 2014, 14). Puhekielisyyden lisäksi omana ulottuvuutenaan voidaan pitää eri murteissa tapahtuvia sanojen alueellisia muunnoksia (Koivisto, 2013, 125).

Sanojen lyhentämisen, puhekielistämisen ja murteiden lisäksi puheessa on myös hienovaraisempia eroja suhteessa ortografiaan. Kirjoitetun kielen muunnosta puheeksi kutsutaan puhutuksi yleiskieleksi, jonka foneemirakenteen katsotaan olevan lähes suorassa suhteessa kirjoitusasuunsa; jokaisella foneemilla on vastingrafeeminsa ja jokaisella grafeemilla oma foneeminsa (Ikola, 2001, 19; Karlsson, 1982, 44). Kuitenkin myös yleispuhekielessä esiintyy ortografiasta poikkeavaa ääntämistä. Tämän tutkielman aiheena on eräs näistä poikkeamista: rajapidennys, joka toteutuu esimerkiksi ilmauksissa [venek:ɑ:tu:] ja [minul:ek:in] sanan- ja liitteenrajaisen [k]:n pitenemisenä. Ilmiön aiheuttavat tietyt, kielenmuutoksen seurauksena nykyään vokaaliloppuiset sananmuotoryhmät, joita mainitaan kieliopissa yleensä toistakymmentä. Rajapidennyksen toteutumisessa on huomattavaa muotoryhmä- ja puhujakohtaista vaihtelua, ja oman lisänsä ilmiön kirjavuuteen tuovat erot eri murrealueiden välillä.

Tässä työssä tarkasteltavaksi huomionkohteeksi on valittu sanan rajalla tapahtuva rajapidennys e-loppuisten nominien muototyypissä, joka on ilmiön aiheuttavista sananmuototyypeistä tunnusomaisin (Nikkinen-Piraccini, 2011; Matušková, 2016). Vaikka kattavaa, murrealueita vertailevaa tutkimusta aiheesta ei ole tehty, tämä on myös eräs muotoryhmistä, joissa on havaittu runsaasti alueellisia eroavaisuuksia (mm. Karlsson & Lehtonen, 1977). Tämän tutkimuksen tavoitteena on selvittää, onko rajapidennys produktiivinen ilmiö: toteutuuko se tutkimusta varten kehitettyjen pseudosanojen yhteydessä. Tätä tutkitaan äidinkielisillä suomenpuhujilla toteutetulla puheentuottokokeella. Koehenkilöiden tuottamien puhunnosten akustisia ominaisuuksia tarkastellaan puheanalyysiohjelma Praatilla (Boersma &

Weenink, 2019), ja tutkimuskysymykset sekä tilastollisin menetelmin tarkasteltavat hypoteesit esitetään aineiston esittelyn jälkeen luvussa 2.3.

Tarkastelun kohteena on puheen ilmiö, joka on yhtäältä muotosidonnainen ja kieliopillinen, toisaalta puhujakohtainen ja osin valinnainen, joten näkemyksiä aiheesta ja tavasta, jolla sitä tulisi nimittää, on useita. Eri lähteissä siitä käytetäänkin lukuisia eri termejä, joita on esitelty luvussa 1.2. Vaikka termin voi nähdä pelkkänä nimenä ilmiölle, ei aihetta määrittävänä attribuuttina, tässä ilmiön foneettiseen toteutumaan keskittyvässä työssä ilmiöstä käytetään sitä konkreettisesti ja tarkimmin kuvaavaa nimitystä *rajapidennys*.

1.1 Rajapidennys

Rajapidennys ilmenee äänteen pidentymisenä, usein kahdentumisena sanojen rajalla. Se on puhutun kielen ilmiö, joka voi toteutua vain yhtäjaksoisessa puheessa edellyttäen, että sanat esiintyvät peräkkäin ilman taukoa. Useimmiten kyseessä on tauon edessä vokaaliloppuisen, tiettyyn muotoryhmään kuuluvan sanan ääntyminen konsonanttiloppuisena siten, että seuraavan sanan alkukonsonantti pidentyy, kuten äännöksessä [tulet:æn:e]. Vokaalialkuista sanaa saattaa puolestaan edeltää pidentynyt glottaalinen sulkeuma, mutta tämä on usein konsonanttikahdentumaa puhujakohtaisempaa. (mm. ISK § 34; Karlsson & Lehtonen, 1977.) Huomion arvoista ilmiössä on myös, että vaikka suomen fonotaksi ei salli sanansisäisiä geminaattoja /hh/, /jj/ ja /vv/, rajapidennys mahdollistaa nämä sanan rajalla. Suomen konsonanttijärjestelmään luetaan määrittelytavasta riippuen kuuluvan 13–17 konsonanttia. Omaperäisiä konsonantteja on 13, ja näistä voivat rajapidentyä kaikki muut paitsi /ŋ/, sillä suomessa ei ole /ŋ/-alkuisia sanoja (Laaksonen & Lieko, 1988, 22).

Sananrajaisten tapausten lisäksi ilmiö toteutuu myös sanan sisällä yhdyssanoissa, kuten sanassa [sadek:atos] sekä liitepartikkeleiden ja tiettyjen adjektiivijohdinten edellä, kuten ilmauksissa [otap:as], [a:vem:ainen] ja [elækel:æinen] (ISK § 34). Iso suomen kielioppi rinnastaa rajapidennyksen yleis- ja erikoisgeminaatioon – toisin kuin Koivisto (2013) tai Mielikäinen ja Palander (2014) – mutta mainitsee, että kyse saattaa olla myös äänteen keston eriasteisesta pidentymisestä, ei välttämättä täydestä kahdentumisesta, johon termi *geminaatio* viittaa.

1.1.1 Määrittelyä

Rajapidennys määritellään kirjallisuudessa hieman eri tavoin. Hakulinen (1979, 49–51) kuvaa sitä sanansisäisenä ja -loppuisena astevaihteluna, Karlsson (1982, 348–349) esittää ilmiön yhtäältä morfofonologisena sääntönä, sillä fonologisesti ei voi ennustaa, mitkä morfeemit sen laukaisevat, ja toisaalta morfofoneettisena sääntönä, sillä pidennysaste on foneettisesti vakiintumaton. Ikola ei määrittele tarkemmin ilmiön olemusta, mutta luonnehtii sen tärkeimpänä suomen ääntämisen ja kirjoituksen välisenä erona (Ikola, 1992, 21; 2001, 20).

Ilmiötä foneettisesti ja kenties perinpohjaisimmin tutkineet Karlsson ja Lehtonen luonnehtivat rajapidennyksen fysikaalisesti todettavaksi konkreettiseksi konsonanttisegmentiksi, ja toisaalta sillä tarkoitetaan heidän mukaansa ”paradigmaattisesti vaihtelevien konsonanttien yhdessä muodostamaa joukkoa, niiden vaihtelukaavaa” (Karlsson & Lehtonen, 1977, 67). He myös kuvaavat ilmiön morfofoneemina, joka esiintyy eräiden sananmuotojen lopussa, ja toisaalta sääntönä ja prosessina tai morfeemipiirteenä (mts. 69–70).

Rajapidennyksellä on siis kirjallisuudessa tarkoitettu hieman eri asioita, mikä on oletettavasti eräs syistä ilmiöstä käytetyn monimuotoisen termistön taustalla.

1.1.2 Merkityseron aiheuttava sandhi

Rajapidennys on eräs suomen sandhi-ilmiöistä. Sandhi on sanojen ja morfeemien rajalla syntyvä äänteellinen ilmiö, kuten sanan tai morfeemin loppuäänteen vahvistuminen tai mukautuminen seuraavan sanan tai morfeemin mukaan. Sandheja on mahdollista jaotella yksityiskohtaisuuden mukaan erilaisiin luokkiin, mutta yleisin jako tehdään sananrajaisiin *ulkoisiin sandheihin* ja sanansisäisillä morfeeminrajoilla tapahtuviin *sisäisiin sandheihin*. Rajapidennyksen voi katsoa kuuluvaksi näistä molempiin tapauksesta riippuen. Termi *sandhi* on alkujaan peräisin sanskritista, ja suomennettuna se tarkoittaa muun muassa yhteen liittämistä. (Itkonen, 1964b, 225; Suomi, 1996, 74; Tieteen termipankki: *Kielitiede:sandhi*.)

Suomen muita sandhi-ilmiöitä on esimerkiksi Länsi-Suomessa tyypillinen sananloppuisen konsonantin elisio, jossa sanan viimeinen konsonantti jää toteutumatta vokaalialkuisen sanan edellä: *tuommoinen otus* ~ [tom:oneotus], *se on iloinen aina* ~ [seoiloneaina]. Tätä yleisempi sandhi on monissa kielissä eri ehdoin vaikuttava nasaaliassimilaatio, joka esiintyy puhutun suomen kaikissa varianteissa: *pojan pallo* ääntyy yhtäjaksoisessa puheessa lähes poikkeuksetta [pojampal:o], *järven muta* puolestaan [jærvem:uta]. Myös yksimorfeeminen

hän toteutuu erilaisena seuraavan sanan alkukonsonantista riippuen: *hän sanoo* ~ [hænsano:], *hän pitää* ~ [hæmpitæ:], *hän fiilaa* ~ [hæmfi:la:], *hän käy* ~ [hæŋkæy]. (Suomi, Toivanen & Ylitalo, 2006, 182; Suomi, 1996, 74–77.)

Muiden suomen sandhien lailla rajapidennys on ortografiassa näkymätön puheen ilmiö. Elisiosta ja nasaaliassimilaatiosta poiketen rajapidennys saattaa tuoda puheeseen kuitenkin merkityseron. Assimiloitunut [pojampal:o] tarkoittaa samaa kuin assimiloitumaton [pojanpal:o], mutta [si:stikoti] ei tarkoita samaa kuin [si:stik:oti]; [k]:n pidentyminen muuttaa toteamuksen käskyksi. Suomessa sananrajaisen konsonantin keston voikin sanoa olevan merkityksiä erottava (Karlsson & Lehtonen, 1977, 26). Tämä kirjoitetun tekstin mahdollisuus reaalistua eri tarkoitteina puheessa sallii myös käyttää kieltä leikkisän kekseliäästi: *Asu tilanteen mukaan, Maista pois, Ei maksa vaivaa ja Saada kiljua sydämen kyllyydestä* saavat hyvin erilaisen merkityksen sananrajaisen konsonantin kestosta riippuen.

Muihin kieliin vertailtaessa esimerkiksi italiassa esiintyy vastaavanlaista konsonantin sanarajanylistä pitenemistä imperatiivissa. *Da!* (suom. *Anna!*) yhdistettynä objektipronomini *mi:*hin (suom. *minulle*) lausutaan puheessa [dam:i]. Erona suomeen – ja yleisesti sandheihin – on kuitenkin se, että puhunnos saa myös ortografiassa muodon *Dammi!* Suomessa käsky kuuluu [an:am:inul:e] kirjoitusasun ollessa *Anna minulle!*

Italiassa esiintyy myös ”oikea” sandhi, alkukonsonantin vahvistuminen, johon suomen rajapidennystä on verrattu (mm. Penttilä, 1969). Ilmiö on rajapidennyksen kaltainen, sillä myös sen laukaisee pidentyvää konsonanttia edeltävä sana yhtäjaksoisessa puheessa. Suomen lailla säännöt ilmiön aiheutumiseen eivät ole täysin yksiselitteiset, sillä sen aiheuttavat muun muassa kaikki painolliset yksitavuiset sanat – ja jotkut painottomista. Rajapidennyksen tavoin ilmiö on myös italiassa jossain määrin puhuja- ja murrekohtainen, vaikka sekin mainitaan kieliopissa eräänä italian puheen piirteenä. (Absalom, Hajek & Stevens, 2004.) Ilmiöstä käytetään yleisimmin syntaktiseen kahdentumiseen viittaavaa nimeä *raddoppiamento sintattico*, mutta termin *rajapidennys* lailla siitä tunnetaan myös nimitys *rafforzamento iniziale*, joka kuvaa ilmiötä *kahdentumisen* sijaan (*alku*)*vahvistumisena*.

1.2 Ilmiön nimityksestä

Rajapidennyksestä on kirjoitettu runsaasti, ja aiheen lingvistisen analysoinnin lisäksi paljon mielipiteitä on herättänyt myös terminologia. Nimityskeskustelua ei tässäkään työssä ole sivuutettu, sillä nimen valinta kertoo osittain suhtautumisesta ilmiöön. On myös kiintoisaa

huomata, kuinka suuren roolin keskusteluissa on saanut ortografiassa näkymätön glottaalinen äänne, joka on osaltaan vaikuttanut tässäkin tutkielmassa ilmiöstä käytettävään nimeen.

Varhaisimpia nimityksiä ilmiölle ovat *aspiraatio*, *final-adspiration* (mm. Becker, 1824; Renvall, 1840), *loppuaspiratsiooni* (mm. Setälä, 1899), sekä oletettavasti näistä suomennettu *loppuhenkonen*. Keskustelua aiheesta on käyty fennistiikassa kyseisillä termeillä tietävästi jo 1700-luvulta alkaen (Nikkinen-Piraccini, 2011). Vuonna 1953 suomen kielen tutkimusta esittelevässä aikakauslehti Virittäjässä julkaistiin kirjoitus, jossa ilmiölle ehdotettiin nimeä *jäännöslopuke* (Ikola, 1953, 431), ja vuonna 1969 sen palstoilla käytiin kiistelystä kyseinen keskustelu, jonka fokus oli lingvistisen määrittelyn sijaan pikemmin pyrkimyksessä löytää ilmiötä parhaiten kuvaava nimitys (Karlsson, 1973, 1–2). Ehdotuksia olivat *jäännöslopukkeen* lisäksi *loppukahdennus* ja *alkukahdennus*. Nimikysymystä käsitteli myös Suomen Akatemian kielilautakunta, mutta se ei ilmaissut kantaansa suositustermistä. (Itkonen, 1969.)

Kaikki vuoden 1969 kirjoittajat olivat yksimielisiä siitä, että termit (*loppu*)*aspiraatio* ja *loppuhenkonen* antoivat ilmiöstä väärän kuvan. Termit tuntuivat liian konkreettisuuden lisäksi harhaanjohtavilta, sillä ne viittasivat lähinnä yhteen ilmiön toteutumaan, glottaaliklusiiliin, joka ei artikulaatiomielessä edes ole ”henkonen”. Glottaaliklusiili on katkoäänen, jossa äänirako suljetaan hetkeksi, jolloin ilman virtaus sen läpi katkeaa. (Karlsson, 1973.) Termejä oli mahdollisesti ylläpitänyt tapa, jolla ilmiötä oli aiemmin kirjallisuudessa merkitty: yleisimmin käytetty apostrofi ’ ja sananloppuinen *h* (mm. Renvall, 1840, 20) olivat saattaneet aiheuttaa harhaluulon, että sanan lopussa esiintyisi jonkinlainen ääntämyksellinen *aspiraatio* (Hakulinen, 1979, 51). *Aspiraatio* viittaakin foneettisena terminä yleensä soinnittomien klusiilien laukeamavaiheen jälkeiseen, [h]-maiseen lisä-äänteeseen, joka toteutuu ennen kuin äänihuulet taas alkavat värähdellä (Karlsson, 1973, 2; Wiik, 1998, 75). Saman termin käyttäminen tarkoittamassa aivan toista ilmiötä olisi harhaanjohtavaa.

Muita ilmiöstä käytettyjä termejä ovat *rajakahdennus* ja *rajageminaatio*, joissa *rajan* käyttäminen termin osana poistaa spekuloinnin siitä, tapahtuuko pidentyminen nimenomaan sanan alussa vai lopussa (Suomi ym., 2006, 184). Vuonna 1969 *rajakahdennus* sai hyväksynnän osalta silloiseen nimityskiistaan osallistuneista, mutta termiä painotettiin myönnettävän käyttämään vain itse ilmiöstä. Ilmiön aiheuttajaa – useissa lähteissä *x-morfeemiksi* kutsuttua segmenttiä (Karlsson, 1982, 349; Laaksonen & Lieko, 1988; Suomi, 1996, 75) – kutsuttaisiin *loppukahdennukseksi*. (Itkonen, 1969.) Myös tämä tukee olettamusta siitä, että termeillä on, kenties tahattomasti, tarkoitettu hieman eri asioita: pidentymistä itseään ja sen aiheuttajaa. Iso

suomen kielioppi käyttää ilmiöstä nimeä *rajageminaatio*, mutta mainitsee, että yleisin siitä käytetty termi on *loppukahdennus* (ISK § 34).

Suomi (1996) ja Suomi, Toivanen ja Ylitalo (2006) edustavat vähemmistöä kutsuessaan ilmiötä nimellä *rajapidennys*. Nimitys on kuitenkin osuva, jos sillä haluaa kuvata ilmiötä mahdollisimman tarkasti, sillä vaikka suurimmassa osassa tapauksia konsonantin kesto kaksinkertaistuu, kyseessä on paikoitellen huomattavasti vähäisempi pidentyminen. Kahdentumiseen viittaavat termit ovat ongelmallisia myös sikäli, että ne olettavat, että kahdentunut segmentti esiintyy kielessä myös yksittäisenä, mikä ei pidä paikkaansa glottaaliklusiilin kohdalla. Esimerkiksi käskyssä [an:aʔ:ol:a] kahdentunutta segmenttiä ei voi nimetä, sillä suomessa ei ole foneemia /ʔ/. Äänne [ʔ] ei myöskään esiinny suomessa yksittäiskonsonantin mittaisena. Jos konsonantin pidentymistä nimitetään kahdentumiseksi, esiintyy myös kysymys, miten syntyneen geminaatan jäsenet sijoittuisivat. Jos vaikkapa [si:stik:oti] katsotaan foneemijonoksi /siistikoti/, sijoittuisiko yksi /k/ sanan loppuun ja toinen toisen alkuun: /siistik#koti/? Vai saisiko *koti* alkuunsa kaksi /k/:ta? Koska foneettisesti pitkää, sananrajaista konsonanttia ei ole fonologisesti mielekästä kuvata kahdella lyhyellä foneemilla, myöskään kahdentumista merkitsevät termit eivät ole tarkoituksenmukaisia. (Suomi, 1996, 77–79; Suomi ym., 2006, 184.)

Suomen ja kollegoiden (2006) näkemyksestä eriävän tulkinnan glottaalisesta äänteestä esittää Lieko (1994) tapaustutkimuksessaan lapsen rajapidennyksen oppimisesta. Glottaaliäänne voi foneettisesti olla joko luja aluke tai glottaaliklusiili. Luja aluke tuotetaan glottaalisella puristussupistumalla, glottaaliklusiili puolestaan glottaalisen sulkeuman avulla. Yksittäinen glottaaliäänne, luja aluke, voi Liekon mukaan esiintyä suomessa minkä tahansa vokaali-alkuisen sanan alussa, mutta ainoastaan rajapidennystapauksissa se voi geminoitua. Hän erottaa glottaaliäänteelle kaksi ominaispituutta ja käyttääkin ilmiöstä termiä *loppukahdennus*. Glottaaliäänteen toteutumiseen eriasteisena viittanevat myös Lennes, Aho, Toivola ja Wahlberg (2006) kirjoittaessaan, että heikosti tuotettua glottaaliklusiilia voi nimittää glottaaliseksi approksimantiksi, puristussupistumaksi. Myös Suomi (2009) mainitsee, että on erotettavissa täysimittainen glottaaliklusiili ja lyhyempi glottalisaatio. Artikkelissaan hän kuitenkin käyttää termejä toistensa synonyymeinä, sillä kyseisessä tutkimuksessa glottaalinen äänne on puhunnoksenalkuinen, jolloin sen pituutta ei voi akustisesti päätellä.

Ilmiötä foneettisesti tutkineet Karlsson ja Lehtonen (1977, 55–65) ovat hekin analysoineet glottaaliklusiilia ja esittäneet huomioita sen fonemisoinnista. He jakavat Suomen ja

kollegoiden (2006) näkemyksen siitä, että voidakseen muodostaa geminaatan, äänteen tulee esiintyä kielessä myös yksifoneemisena. Liekon (1994) esittämää mahdollisuutta glottaali-äänteen postuloimisesta kaikkien vokaalialkuisten sanojen alkuun he sen sijaan eivät pidä morfofonologisesti hyväksyttävänä. Fonemisaatiot, kuten /ʔupota/ tai /ʔovi/ eivät heidän mukaansa vastaa suomen puhujan kielitajua, joten koska yksittäisäännettä ei ole, myöskään geminaatiosta ei voi puhua. He myös huomauttavat, ettei oppositiota /ʔ/:n ja /ʔʔ/:n välillä ole olemassa. Ilmeisestä glottaaliäänteen geminaation vastustuksestaan huolimatta he käyttävät ilmiöstä nimitystä *alkukahdennus*.

1.3 Rajapidennyksen aiheutuminen

Rajapidennys on eräs suomen ominaispiirteistä, johon äidinkielen puhuja ei juuri kiinnitä puheessaan huomiota, mutta jonka vieraskielinen suomenoppija joutuu erikseen opettelemaan. Opettelusta tekee haastavaa moniin kielioppiseikkoihin nähden se, että yksiselitteistä sääntöä sen tuottamisesta ei ole olemassa, ja koska kyse on puheen ilmiöstä, myös puhujakohtaista vaihtelua on paljon. Rajapidennyksen aiheutuminen voidaan kuitenkin esittää kolmen säännön avulla (Laaksonen & Lieko, 1988, 22–24; Verkkokielioppi):

Sääntö 1. Seuraavan sanan alkukonsonantti kahdentuu edellisen sanan loppuun. Rajapidennys toteutuu sinä konsonanttina, jolla seuraava sana tai liite alkaa: [kirjek:uori], [minul:ek:in].

Sääntö 2. Vokaalialkuisen sanan edellä rajapidennys voi toteutua lujana alukkeena, glottaaliksiilina – tai jäädä toteutumatta: [an:aʔ:ol:a], [meneʔ:ulos].

Sääntö 3. Tauon edellä rajapidennystä ei tapahdu.

1.3.1 Rajapidennyksen aiheuttavat muotoryhmät

Ilmiötä määrittelevien sääntöjen vaatimus yhtäjaksoisesta puheesta ei ole ainoa ehto rajapidennyksen reaalistumiselle, sillä sen lisäksi ilmiö edellyttää tietyn sananmuototyypin toteutuakseen. Sananmuotoryhmiä, joiden yhteydessä ilmiö yleiskielessä toteutuu, mainitaan kieliopissa yleensä toistakymmentä. Taulukossa 1 on esiteltynä niistä tyypillisimmät ottamatta kantaa yleisyyteen; ensimmäisenä oleminen ei indikoi, että rajapidennys olisi kyseisen muotoryhmän yhteydessä yleisin tai viimeisen kohdalla harvinaisin, sillä eri lähteet määrittävät esiintyvyyden hieman eri tavoin. Ensimmäiset neljä muotoryhmää ovat kuitenkin ne, joissa rajapidennys toteutuu systemaattisimmin ja vahvimmin (Karlsson, 1982, 349–350).

Taulukko 1. Rajapidennyksen aiheuttavia sananmuototyyppejä. Mukailtu Ison suomen kieliopin (ISK § 34), Laaksosen ja Liekon (1988) ja Verkkokieliopin mukaan.

Muototyyppi	Esimerkki
Yks. 2.pers. imperatiivi	[tulet:æn:e], [an:aʔ:ol:a]
A-infinitiivi	[sa:kotul:as:in:e], [voimen:æʔ:elokuvi:n]
Verbin kieltomuodot (ei konditionaali, eikä älä/älkää -sanojen yhteydessä)	[en:otak:ahvia], [ælkæ:tulkol:æhel:e], [etot:anem:aitoa], [eiuskotav:alheita]
<i>Vene</i> -tyypin nominien yks. nominatiivi	[puoluet:uki], [hernek:eit:o], [hamer:ypisty:], [tervet:uloa]
Allatiivi	[meil:et:ule:vieraita], [ihol:eʔ:aurinkoa]
3.pers. possessiivisuffiksi <i>-nsä</i>	[vaimonsak:ans:a], [isiensæl:uval:a]
<i>-sti</i> -adverbit	[runsa:stiv:et:æ], [hyvæstis:it:en]
<i>-nne</i> -adverbit	[sin:et:æn:et:uon:e]
<i>(i)tse</i> -adverbit (prolatiivi)	[sæhkeitsej:apuhelimitsen:euvoit:i:n]
Attribuuttina oleva komitatiivi	[raskainek:as:eine:n]
Yksittäisiä sanoja	[ki:n:iv:eti], [itser:akas], [kolmek:erta:]
Epämuodollisia puhekielisiä ilmauksia	[en:o:m:en:ys:in:e], [seontul:ut:akasin], [tulilaske:k:as:an], [kuol:um:ies]

1.3.2 Kielihistoriallista taustaa

Rajapidennys ilmiönä juontuu kielitieteen mukaan kielihistoriasta: sanat, joiden yhteydessä rajapidennys tapahtuu, ovat nykyään vokaaliloppuisia, mutta päättyivät kantasuomessa konsonanttiin (mm. Hakulinen, 1979; Itkonen, 1964a; Verkkokielioppi). Taulukon 1 muotoryhmistä näitä ovat kaikki paitsi puhekieliset ilmaukset – joista useimmat ortografiassa edelleen päättyvät konsonanttiin: *kuollut mies* ~ [kuol:um:ies]. E-loppuisten, sanan *vene* lailla taipuvien nominien ryhmä päättyi alkujaan [k]:hon tai [h]:hon: *venek, pistek ja herneh, ahneh*. Muita [k]-päätteisiä muotoryhmiä olivat muun muassa A-infinitiivi, verbien kieltomuodot, prolatiivi, attribuuttina oleva komitatiivi ja yksikön toisen persoonan imperatiivi. Allatiivi päättyi [n]:ään ja [t] oli oletettavasti monikon nominatiivipääte, kuten sanassa *kolmet*. (Hakulinen, 1979, 49–52; Laaksonen & Lieko, 1988, 22.) Puheen tutkimiselle ominainen

haaste ilmenee kuitenkin siinä, että kantasuomesta ei ole olemassa puhenäytteitä, joten sananloppuisten konsonanttien laadusta ei ole täyttä varmuutta (Itkonen, 1964a, 18).

Rajapidennystä ja sananloppuisia konsonantteja on tutkittu runsaasti – erityisen perinpohjaisesti Itkonen (1964a) – mutta karkeasti yksinkertaistaen voitaneen sanoa, että rajapidennyksen synnyssä lieene alkujaan kyse loppukonsonantin assimilaatiosta. Esimerkiksi käsky *tulek pois* on assimiloitunut ensin muotoon *tulep pois*, ja lopulta myös sananloppuinen [p] on kadonnut jättäen jälkeensä astevaihtelun. On myös mahdollista, että sananloppuinen konsonantti ei ole kokonaan kadonnut, vaan sen edustajaksi on jäänyt glottaalinen äänne, joka reaalistuu rajapidennystapauksissa. (Häkkinen, 1994, 161–165.) Tätä kadonnutta äännettä on nimitetty *x-morfeemiksi*, joka toteutuu eri äänneympäristöissä erilaisina konsonantteina, kuten [venekɨ:a:tui], [venelɨ:asket:i:n]. *X-morfeemi* -nimitys juontuu rivinylisestä x-kirjaimesta, jolla sitä on kielitieteessä merkitty. (Ikola, 2001, 20–21; Karlsson & Lehtonen, 1977; Laaksonen & Lieko, 1988; Verkkokielioppi.)

Myös aiemmin mainitussa rajapidennykseen rinnastettavissa olevassa italian alkukonsonantin vahvistumisessa on kyse samasta ilmiöstä, taannehtivasta assimilaatiosta tiettyjen, alkujaan konsonanttipäätteisten sanojen yhteydessä, joiden loppukonsonantti on ajan kuluessa redusoitunut pois. Esimerkiksi *ad* on muuttunut muotoon *a*, *tres* muotoon *tre*. Lähteissä myös mainitaan, että todellisuudessa tämä konsonantin elisio reaalistuu vain ortografiassa, sillä puheessa konsonantti ei ole kadonnut, vaan se yhdistyy seuraavan sanan alkukonsonanttiin vahvistaen sitä – kuten pääasiassa suomessakin. *Vado a casa* (suom. *menen kotiin*) toteutuu puheessa muodossa [va:doak:a:sa]. (Treccani; Penttilä, 1969.)

1.4 Rajapidennyksen toteutuminen ja aluekohtaisuus

Rajapidennyksen toteutuminen ei kuitenkaan riipu ainoastaan säännöistä ja muotoryhmästä. Ne ovat edellytys pidennykselle, mutta se, reaalistuuko rajapidennys puheessa vai ei, vaihtelee puhuja- ja murrekohtaisesti. Huomattavaa on siis, että vaikka rajapidennys edellyttää jonkin muotoryhmistä toteutuakseen, muotoryhmä ei edellytä rajapidennystä – ainakaan suurimmassa osassa tapauksia. Vaihtelua ja epäsäännöllisyyttä saattaa olla myös samalla puhujalla eri muotoryhmissä tai eri äänneissä (mm. Karlsson & Lehtonen, 1977).

Rajapidennys ilmiönä kuuluu useimpiin suomen murteisiin (Lieko, 1994, 3; Lehtimäki, 2012), mutta ilmiössä on myös poikkeuksia eri murrealueilla. Kielihistoria selittää aluekohtaisten erojen johtuvan aiemmin konsonanttiloppuisten sanojen loppukonsonantin kehityksestä eri

suuntaan ja eri aikaan eri alueilla (Häkkinen, 1994, 161–165). Lounaismurteissa rajapidennys on koskenut alun perin vain soinnittomia klusiileja ja [s]:ää, ja esimerkiksi kaakkois-, lounais- ja kaakkoishämäläisissä murteissa rajapidennys toteutuu nykyäänkin lähinnä vain klusiili-alkuisen liitepartikkelin edellä, ei niinkään sanan rajalla (ISK § 35; Verkkokielioppi). Kaksikielisten raja-alueiden tuntumassa rajapidennystä esiintyy vähemmän, sillä ilmiötä ei ole ruotsissa tai venäjässä (Lieko 1994, 2).

Rajapidennyksestä ei ole tehty kattavaa, murrealueita vertailevaa tutkimusta, mutta Karlssonin ja Lehtosen tutkimuksen (1977) voinee katsoa jossain määrin sellaiseksi, sillä vertailtavana on eri rajapidennysmuotoryhmien lisäksi kolme erimurteista aluetta, Turku, Kuopio ja Haapamäki. Tutkimuksen perusteella Turku eroaa verrokeistaan vähäisemmällä pidentämisellä, ja tämä ilmenee erityisesti e-loppuisten nominien muotoryhmässä. Yksikön toisen persoonan imperatiivin yhteydessä rajapidennys puolestaan toteutuu lähes yhtäläisenä kaikilla paikkakunnilla. Tietyt ilmaukset, kuten [tæn:ep:æin], tuotetaan pidentyneinä kaikilla kolmella alueella, minkä voi tulkita jossain määrin puhunnosten leksikaalistumiseksi; puhunnos tuotetaan ja hahmotetaan yhtenä kokonaisuutena. Murrealue-erojen lisäksi tutkimuksen keskeisiä tuloksia on puhujakohtaisen vaihtelun oleminen suurta samankin murre- ja muotoryhmän sisällä. (Karlsson & Lehtonen, 1977.)

1.5 E-loppuiset nominit: [hernek:eit:o] ja [nuk:ekoti]

Osassa muotoryhmiä rajapidennys tuo puhunnokseen merkityseron. Näin on esimerkiksi yksikön toisen persoonan imperatiivissa; käsky kuuluu käskynä, kun pidennys tuotetaan verbin ja sitä seuraavan sanan rajalle: [sois:anakultainen]. Toisissa tapauksissa, kuten *vene*-tyypin sanoissa, pidentäminen ei muuta merkitystä, [hernek:eit:o] ja [ahnem:ies] merkitsevät samaa rajapidennyksellä ja ilman. Tämän nykyään e-loppuisten nominien muotoryhmän juuret ovat vanhoissa sanoissa, mutta sanatyypin on produktiivinen, ja sanoja muodostetaan ja lisätään ryhmään jatkuvasti. Nykysuomen sanakirja mainitsee noin 1200 tämän sanatyypin nominia, mutta todellisuudessa niitä on enemmän, ja ryhmä kasvaa koko ajan (Verkkokielioppi).

Toinen e-loppuinen nominiryhmä ovat sanan *nalle* lailla taipuvat sanat, joissa rajapidennystä ei tapahdu. *Nalle*-tyypin nomineita on suomessa huomattavasti *vene*-tyypin sanoja vähemmän – Nykysuomen sanakirja mainitsee niitä olevan alle sata (Verkkokielioppi) – eikä ryhmä ole erityisen produktiivinen, se kasvaa lähinnä lainasanojen, kuten *single* ja *college*, myötä.

Lisäksi on olemassa poikkeuksia, kuten taivutusparadigmaltaan *nalle*-ryhmään kuuluvat sanat

kolme ja *itse*, joiden nominatiivin yhteydessä rajapidennys kuitenkin voi toteutua, sekä sana *viime*, jonka jälkeen käytetään niin ikään pidentynyttä ja pidentymätöntä muotoa (Ikola, 2001, 21; ISK § 34; Laaksonen & Lieko, 1988, 23–24; Verkkokielioppi).

Vaikka *nalle*-sanoja on huomattavasti *vene*-sanoja vähemmän, rajapidennyksen toteutumisen ennustamattomuus e-nomineissa lienee suomea opettelevalle haastavaa, sillä varsinaista sääntöä ei ole, vaan sanat tulee osata sijoittaa oikeaan ryhmään. E-loppuisten nominien ääntymisen epä johdonmukaisuus käy ilmi myös lapsen rajapidennyksen oppimisessa: muihin muotoryhmiin – erityisesti toisen persoonan imperatiiviin, kielteiseen preesensiin ja allatiiviin – verrattuna vähäinen ja vakiintumaton pidentäminen e-nomineissa selittyy *nalle*-tyypin sanojen aiheuttamalla hämmennyksellä (Lieko, 1994). Tämä on ymmärrettävää, sillä lapsen puheessa rajapidennystä aiheuttamattomat *nukke* ja *nalle* esiintyvät usein verrattuna aikuisen sanastoon, jossa *vene*-tyypin nominit puolestaan dominoivat. Myös Räisänen mainitsee, että hänen tutkimustaan varten tarkkailemiensa lasten puheessa sana *nalle* oli frekvenssiltään ylivoimainen kaikkiin *vene*-tyypin nomineihin nähden. Tämä näkyi myös *vene*-tyypin sanojen taivuttamisessa virheellisesti *nalle*-sanan paradigman mukaan: ”kaksi *partakonea* isillä” (kaksi *partakonetta* isillä), ”siellä oli (*lento*)*koneja*” (siellä oli (*lento*)*koneita*). (Räisänen, 1975, 256–257.)

E-loppuisten nominien pidentyminen tai pidentymättä jääminen on rajapidennyksestä myös kenties yleisesti tiedostetuin tapaus. Tämä ilmenee muun muassa kyselytutkimuksesta pirkanmaalaisille ja uusimaalaisille maallikkokielenkäyttäjille, joiden yleisimmin rajapidentämä muotoryhmä on *vene*-tyypin nominit, erityisesti niiden esiintyessä yhdyssanan alkuosana. Toiseksi yleisin muototyyppi on yksikön toisen persoonan imperatiivi, ja kolmantena on sana *itse*. Tutkimuksen mukaan ilmiön puuttuminen huomataan harvemmin kuin sen olemassaolo, mutta jos puuttuminen huomataan, se koetaan hassuksi tai oudoksi. (Matusková, 2016.)

Myös maallikkokielenkäyttäjien verkkokeskusteluihin pohjautuvan tutkimuksen mukaan kuuliija saattaa yhdistää rajapidennyksen puuttumisen eri murteisiin sekä ei-äidinkielisiin puhujiin, mutta myös hienostelemaan, teeskentelevään ja epäluonnolliseen puheeseen. Tyypin-esimerkkinä mainitaan rajapidentymättömyys e-loppuisten nominien yhteydessä. *Vene*-tyypin sanat on mainittu keskusteluissa lähes kolme kertaa niin usein kuin toiseksi tyypillisin rajapidennyksen aiheuttaja, yksikön toisen persoonan imperatiivi, ja 13-kertaisesti verrattuna kolmannella sijalla oleviin sanoihin *itse*, *kolme* ja *viime*. Erityisesti rajapidennyksen

toteutumista odotetaan yhdyssanan alkuosana esiintyvän e-nominin yhteydessä. (Nikkinen-Piraccini, 2011.)

Verkkokeskusteluista nousee esiin myös lekseemin *hernekeitto* mahdollinen vakiintuminen edustamaan koko rajapidennysilmiötä maallikoiden keskusteluissa, siinä määrin useasti kyseinen sana mainitaan aineistossa. Tämä lienee verrattavissa maallikoiden murteista käyttämää metakieltä tutkineen Mielikäisen havaintoihin; tietyt ilmaukset tai käsitteet ovat vakiinnuttaneet paikkansa ei-kielitieteilijöiden puhuessa kielestä. Tällaisia ovat esimerkiksi *tois pual jokke* - ja *ookko nää (Oulusta)* -imitaatiot, jotka yhdistetään poikkeuksesta tiettyihin murteisiin. (Nikkinen-Piraccini, 2011, 65; Mielikäinen, 2005, 101–102.) Samaan tapaan sana *hernekeitto* antaa rajapidennysilmiölle kansantajuksen reaalistuman.

Rajapidennyksen toteutumista sen yleisimmin aiheuttavien muotoryhmien yhteydessä on tutkittu myös kyselytestimenetelmällä (Lehtimäki, 2012). Morfofonologisessa tutkimuksessa e-loppuisten nominien ryhmään oli lisätty myös epäsanvoja, joiden avulla tarkasteltiin ilmiön produktiivisuutta. Informanttien tehtävänä oli koelauseiden luvun lisäksi taivuttaa e-loppuisia epäsanvoja, millä mitattiin, sijoittiko lukija sanat *vene-* vai *nalle-*ryhmään. Ilmiön toteutuminen osoittautui vaihtelevaksi koehenkilöiden ja erilaisten epäsanvojen välillä. Osa haastatelluista ei kokenut mielekkääksi lisätä mitään sanaan, joka on itselle vieras, mutta tuottaisi rajapidennyksen sanan ollessa oikea. Ilmiön voi tutkimuksen perusteella kuitenkin katsoa olevan jossain määrin produktiivinen puheen piirre.

1.6 Suomen kvantiteetti

Puhe koostuu äänneistä, joiden kestot vaihtelevat eri kielissä eri syistä. Ilmeisin syistä on puhenopeus; mitä nopeammin sana sanotaan, sitä lyhyempiä äännekestot ovat. Äänneiden kestot eivät ole absoluuttisia, vaan suhteellisia toisiinsa nähden, jolloin on mahdollista erottaa toisistaan kestoiltaan pitkiä ja lyhyitä äänneitä. Äännekestoja mitataan sekunteina ja millisekunteina, ja näitä mitattuja kestoja kutsutaan kvantiteetiksi. (Suomi ym., 2006, 118–119.)

Suomi on kvantiteettikieli, jossa kvantiteettioppositio toteutuu merkitystä erottavana niin vokaaleissa kuin konsonanteissa ja sekä painollisissa että painottomissa tavuissa. Äänneille erotetaan kaksi kestoastetta, lyhyt ja pitkä. Lyhyttä äännettä merkitään yhdellä grafeemilla ja pitkä kahdella, vaikka kyseessä on yksi yhtenäinen pitkä äänne. Kvantiteetin toteutuminen on riippumaton siitä ympäröivistä äänneistä tai asemasta sanassa, ja suomessa esiintyykin kaikkia mahdollisia kvantiteettiyhdistelmiä, lukuun ottamatta fonotaksin vastaisia pitkiä

konsonantteja sanan alussa tai lopussa. Pitkiä kvantiteettiasteita esiintyy suomessa verrattain paljon; tällä tavoin on mahdollista korvata foneemien vähyyttä (Karlsson, 1969, 355; Nahkola, 1987, 11). Vaikka tonaalisilla vihjeillä, kuten perustaajuuden muutoksella, on havaittu olevan vaikutusta äänteen havaitsemiseen pitkänä tai lyhyenä (Vainio, Järvikivi, Aalto & Suni, 2010), pääasiallinen pitkä–lyhyt-erottelu perustuu äänteen suhteelliseen keston. Distinktiivisiä kestoeroja maksimaalisesti hyödyntävän suomen rajapidennystä tutkittaessa sivutaankin samalla kvantiteettia. Myös Lieko mainitsee, että osatakseen tuottaa rajapidennyksen lapsen tulee pystyä tuottamaan pitkän ja lyhyen kvantiteetin oppositio (Lieko, 1994, 21).

Maailman kieliin suhteutettuna on poikkeuksellista, että suomessa esiintyy enemmän pitkiä konsonantteja kuin pitkiä vokaaleja, sillä yleinen tendenssi kvantiteettikielissä on päinvastainen. Konsonanttien kontrastiivinen pituusoppositio on maailmanlaajuisesti tyypillisempää obstruenteilla kuin resonanteilla, ja resonanteilla ääntöväylän avoimuus vaikuttaa siten, että sonoriset ovat epätyypillisemmin kontrastiivisen pituusopposition edustajia. Niinpä esimerkiksi klusiilit ovat tyypillisempiä kuin nasaalit ja nasaalit puolestaan tyypillisempiä kuin likvidat. Tämänkin suhteen maailman kielissä esiintyy kuitenkin poikkeuksia. Eräänä sellaisena voitaneen pitää suomea, jossa esiintyy verrattain vähän pitkiä soinnittomia klusiileja, mutta runsaasti pitkiä sonorantteja ja frikatiiveja. Tämä selittynee osittain suomen morfologialla, esimerkiksi verbien aikamuodoilla: *olen noussut, on tullut*. (Gordon, 2016, 51–57.) Suomea analysoitaessa huomioon ei kuitenkaan liene otettu rajapidennystapauksissa toteutuvien pitkien konsonanttien laatua tai eri murteissa tapahtuvia äänteiden pidentymisiä.

1.7 Paino

Keston lisäksi toinen äänteen havaittuun ja mitattuun pituuteen vaikuttava ominaisuus on paino. Luonnollisessa puheessa painotusta tapahtuu sekä lause- että sanatasolla. Lausepainolla tarkoitetaan tietyn sanan tai tiettyjen sanojen painottamista lauseessa, ja sen pääasiallinen tarkoitus on korostaa asiasisältöä, kiinnittää huomio tiettyyn osaan lausetta. Sanapainolla puolestaan tarkoitetaan sanan eri tavujen painottamista. Sanapainosäännöt vaihtelevat kielikohtaisesti, ja sanapainoa käytetään yleisimmin ilmaisemaan kieliopillisia merkityksiä. Esimerkiksi englannissa sillä erotetaan sanaluokkia [*ˈɪmpɔːt*] (suom. *maahantuonti*) ja [*ɪmˈpɔːt*] (suom. *tuoda maahan*). (Wiik, 1988, 107–111.)

Suomessa voidaan erottaa kolme tavujen keskinäisiä suhteita kuvastavan sanapainon astetta: pääpainollinen, sivupainollinen ja painoton (Suomi ym., 2006, 220). Pääpaino lankeaa

suomessa aina sanan ensimmäiselle tavulle, kun taas sivupainon paikka on sidoksissa tavujen segmentaaliseen rakenteeseen (Karlsson, 1982, 150–151). Maailman kielissä kiinteän sanapainon toteutumista yleisin on toiseksi viimeiselle tavulle lankeava paino ja toiseksi yleisin suomen kaltainen, sanan ensitavuun sidonnainen paino, joka esiintyy 18,3%:ssa kieliä (WALS). Kiinteän sanapainon kielissä painolla on yleisinhimillisen rytmillisyyteen pyrkimyksen lisäksi tehtävänsä sananrajan osoittimena: suomen puhuja esimerkiksi sijoittaa automaattisesti pääpainollisen tavun sanan alkuun (Wiik, 1998, 109).

Painon merkitys on rajapidennyksenkin kannalta ratkaiseva, sillä rajapidennys toteutuu sanojen rajalla – jolloin painollisella ensitavulla on osuutensa ilmiön synnyssä. Rajapidennys toteutuu myös puhumaan opettelevan lapsen puheessa sitä varmemmin, mitä painollisempi sana on kyseessä (Lieko, 1994, 16). Monien muiden kielten tapaan suomenkin painollisessa asemassa olevat segmentit ja tavut ovat foneettisesti pisimmillään (Nahkola, 1987, 15; Suomi, ym., 2006, 223), ja joissain tapauksissa painollinen lyhyt vokaali on mitattu kestoaltaan jopa pidemmäksi kuin painoton pitkä vokaali. Vastaavaa on havaittu myös konsonanteissa: esimerkiksi lausekkeen *tarkka kirurgi* [k]-äänteet osoittautuivat kestoiltaan lähes identtisiksi, vaikka ensimmäisessä on kyseessä pitkä kvantiteettiaste, toisessa lyhyt (Vainio, 2001, 23–26). Teoreettisesti yksinkertainen kesto-oppositio ei näin ollen foneettisena toteutumana ole aivan yksiselitteinen. Myös Lieko (1994) mainitsee vaikeuden erottaa geminaatta ja yksittäis-konsonantti toisistaan niiden ollessa painottomassa asemassa.

1.8 Äänteiden kesto ja laatu

Suomessa on kaksi äännekestokategoriaa, mutta äänteiden toteutumia on lukematon määrä, ja ne vaihtelevat jopa samalla puhujalla. Tutkimuksissa on myös havaittu ero luku- ja spontaani-puheen välillä: puhe- ja artikulaationopeus ovat suomessa tavallisesti suurempia lukupuheessa kuin luonnollisessa puhunnassa (Toivola, 2011, 39–40). Puhenoitusmittaukset tehdään tavutasolla, mutta vaikutus heijastuu myös yksittäisiin äänteisiin. Suomen sanafonologisen opposition yksittäisäänteen ja geminaatan välillä nähdään kuitenkin toteutuvan foneettisesti lähes ihanteellisella tavalla, sillä äännekestojen suhde on 1:2 (Karlsson & Lehtonen, 1977). Tämä pätee erityisesti konsonantteihin, mutta osittain myös vokaalisegmentteihin (mts. 59).

Kvantiteettikielille ominaisten lingvistiksi tarkoitettujen suhteellisten kestojen sekä puhujakohtaisen kestovaihtelun lisäksi äänteiden kestoihin vaikuttavat myös niiden ääntämiseen vaikuttavat fysiologiset ominaisuudet. Näitä ovat esimerkiksi konsonanttien

artikulaatiopaikka tai -tapa ja vokaalien etisyys tai suppeus. Myös äänteen ympäristöllä on vaikutuksensa, sillä mitä pidempi on artikulaattorien liikerata äänteestä toiseen, sen pidemmäksi muodostuu myös äänteen kesto. (Lehiste, 1970, 18–20.)

Suomen äänteille on mitattu niille ominaisia pituuskestoja suhteuttamalla tietyn äänteen kesto kaikkien äänteiden keskimääräiseen keston. Rajapidennyksen kannalta kiinnostavimpia ovat konsonantit, joiden Lehtosen vuonna 1970 mittaamat ominaiskestot Nahkola esittelee väitöskirjassaan (1987, 13–14). Äännekestot ovat huomattavan erilaisia, esimerkiksi [p] on kaksinkertainen [d]:hen nähden, ja likvidat alle puolet soinnittomien klusiilien tai [s]:n pituudesta. Geminaattojen ominaiskestot noudattelevat yksittäiskonsonanttien kaavaa: obstruentit – erityisesti soinnittomat – ovat pidempiä kuin resonantit. Suhteellinen ero geminaatan ja yksittäiskonsonantin välillä on kuitenkin obstruenteilla pienempi kuin resonanteilla; klusiili- ja frikatiivigeminaatat ovat noin kaksi kertaa yksittäiskonsonanttiansa pituisia, likvidat ja nasaalit 2,5-kertaisia yksittäisiin nähden. Geminaatoissa on mitattuna myös sanansisäinen [ŋ:], mutta ei rajapidennyksessä esiintyviä [h:], [j:] tai [v:]. Nahkola mainitsee, että eri tutkimukset ovat antaneet konsonanttien kestoista hieman eriäviä tuloksia, mutta soinnittomat ovat osoittautuneet poikkeuksetta soinnillisia pidemmiksi (mts. 13).

1.8.1 Äänteiden laatu ja rajapidennys

Konsonantin laatu vaikuttaa jossain määrin myös rajapidennykseen. Karlssonin ja Lehtosen (1977) mukaan suomen fonologiseen perusjärjestelmään kuuluvat konsonantit pitenevät useammin kuin esimerkiksi [f] tai [b], ja obstruentit pitenevät resonantteja useammin. Soinnittomien obstruenttien lähtökohtaisesti pidempi kesto ja suurempi erottuvuus puheesta saattavat olla syynä myös siihen, että lapsi tuottaa rajapidennyksen huomattavasti useammin obstruenteissa kuin resonanteissa (Lieko, 1994). Toinen syy niiden suurempaan pidentymiseen saattaa olla se, että lapsen on helpompaa tuottaa nämä ytimekkäät äänteet. Lieko vertaa tutkimuksessaan tapauksia, joissa allatiivia seuraa konsonanttialkuinen sana. Rajapidennys toteutuu obstruenteissa [k, p, t, s, h] 64 prosentissa ja resonanteissa [m, n, l, v, j, r] 38 prosentissa potentiaalisia tapauksia.

Myös maallikoiden verkkokeskusteluihin pohjautuvan tutkimuksen tuloksista käy ilmi, että tyypillisinä pidetyissä rajapidennystapauksissa pidentyvänä konsonanttina on yleisimmin obstruentti, erityisesti soinniton klusiili (Nikkinen-Piraccini, 2011, 65).

1.8.2 Tapaus [h]

Poikkeus obstruenteista on [h]. Sen harvinaisuus geminaattana mainitaan foneettisena universaalina, ja syyksi tähän esitetään sen muihin äänteisiin verrattuna suuri ilmavirran kuluttavuus (Itkonen, 1987, 198). Myös lapselle sen pidentämisen oppiminen on vaikeampaa kuin muiden – erityisesti suun etuosassa tuotettavien (Räisänen, 1975, 255) – konsonanttien, ja sitä perustellaan äänten artikulaatiopaikan vaihtelevuudella (Lieko, 1994). Suomen [h]:lla ei ole omaa ääntymäpaikkaa (Raimo & Ojala, 2009, 181), ja sillä on monta allofonia, jotka vaihtelevat äännekontekstin mukaan. Näitä ovat muun muassa [h ç x]. Useimmiten [h] on soinnillinen, mutta se voi olla myös soinniton esiintyessään soinnittomien äänteiden yhteydessä, kuten sanassa *pihka*. Mahdollisesti vaihtelevuudestaan johtuen sen ääntämisen oppimisen on huomattu olevan haastavaa myös S2-oppijoille (Aho, Toivola, Karlsson & Lennes, 2016).

Ominaisuuksiensa erityislaatuisuuden lisäksi [h]:n on havaittu käyttäytyvän poikkeavasti rajapidennystapauksissa. Karlssonin ja Lehtosen foneettisessa tutkimuksessa (1977) se piteni harvemmin kuin muut konsonantit, ja rajapidennyksen vaihtelu [h]:ssa mainitaan myös Isossa suomen kielioipissa (ISK § 34). Myös Ikola huomauttaa, ettei rajapidennystä [h]:n edellä edellytetä: on yhtä oikein sanoa [paneħ:at:upæ:hæsi] kuin [paneħat:upæ:hæsi] (Ikola, 2001, 21). Suomen fonotaksin mukainen esiintymättömyys sanansisäisenä geminaattana saattanee vaikuttaa asiaan jossain määrin, vaikka [v]:n ja [j]:n vastaavasta vähemmästä pitenemisestä tai rajapidennyksen vaihtelevuudesta ei ole mainintoja. On esitetty myös, että rajapidennyksen yhteydessä tulisi analysoida [h]:n pituuskeston sijaan sen kvalitatiivista soinnillisuus–soinnittomuus-oppositiota (Penttilä, 1963, viitattu lähteessä Itkonen, 1964a, 57). Vaikka ehdotus ei ole saanut kannatusta, se kertoo silti jotakin äänten erikoispiirteisyydestä.

Äänten erityisyydestä kertoo osaltaan myös se, että varhaiskantasuomi ei tuntenut sitä, vaan [h] kehittyi hiljalleen [š]-äänteestä (Itkonen, 1964a, 17). Sen ”levottomuudesta” konsonanttina kielii niin ikään metateesien runsaus, ennakoiva artikulaatio: *veneħtä ~ venħettä, unoħtaa ~ unħottaa, ereħtyä ~ erħettyä*. (Rapola, 1966, 245–251.) Osa metateesin sisältävistä sanoista on edelleen käytössä, esimerkiksi *valħetella* on vanhahtava, mutta hyvin ymmärrettävä.

2 Tutkimusmenetelmät ja aineisto

Suomessa on kahdenlaisia e-loppuisia nomineita: rajapidennyksen aiheuttavia *vene*-tyypin sanoja ja *nalle*-tyypin sanoja, joiden yhteydessä rajapidennys ei toteudu. *Vene*-muotoryhmän on todettu olevan runsas ja produktiivinen, toisin kuin *nalle*-ryhmän. Kiinnostavaa onkin, kumpaan ryhmään äidinkielen suomenpuhujia sijoittaa e-loppuisen pseudosubstantiivin tai nimen lukemassaan lauseessa. Tätä selvitetään puheentuottokokeella kerätyllä aineistolla.

Aineisto koostuu 14 koehenkilöltä äänitetyistä tutkimusta varten kehitetyistä lauseista. Äänitettyjä lauseita on yhteensä 2630, joista analysoitavia koelauseita on 2412. Kaikkiaan tutkimuksessa esiintyy neljää ilmiön mahdollisesti laukaisevaa herätesanatyyppeä. Vokaali–konsonantti-rakenteeltaan sanat ovat kahdenlaisia, ja varsinaisena huomionkohteena, eli pidentyvänä tai pidentymättä jäävänä ääntenä on kolme erilaista konsonanttia.

2.1 Koehenkilöt

Koehenkilöitä on 14, naisia kahdeksan ja miehiä kuusi. Informanttien ikäjakauma on 25–58 vuotta mediaani-ikä ollessa 38 vuotta ja keski-ikä 41 vuotta. Koehenkilöt ovat pääosin pääkaupunkiseudulta, mutta mukana on myös Lohjalla asuva ja Nokialla asuva henkilö. Yksi puhujista on alkujaan kotoisin Oulusta, toinen Kotkasta ja kolmas Kokkolasta, mutta he ovat asuneet Helsingissä jo kymmeniä vuosia. Kahden syntymäpaikasta Helsingistä ja aina siellä asuneen informantin vanhemmat ovat kotoisin muualta kuin pääkaupunkiseudulta.

Kukaan koehenkilöistä ei ole opiskellut kielitieteitä, suomea tai muitakaan kieliaineita ylemmällä kuin lukiotasolla, eli kaikki ovat äidinkielistä maallikkokielenkäyttäjiä. Kaikilla on peruskoulun jälkeisiä opintoja, kuten ylioppilas-, ammattikoulu- tai alempi korkeakoulu-tutkinto, mutta ei ylempää korkeakoulututkintoa. Kaikki ovat työssäkäyviä henkilöitä, jotka käyttävät puhetta vaihtelevassa määrin työtehtävissään. Ammattien edustus on laaja, koehenkilönä on muun muassa tanssinopettaja, radiojuontaja, logistiikkatyöntekijä, myyjä, lastentarhanopettaja, laulaja, talousassistentti ja kirjaston esimies. Kaikki informantit ovat oman näkemyksensä mukaan vähintään jossain määrin tietoisia tavastaan puhua – valtaosa erittäin tietoisia. Kaikki paitsi yksi valitsivat myös taustatietolomakkeen vaihtoehdon ”kyllä” kysymykseen huomion kiinnittämisestä muiden tapaan puhua.

Kaikki osallistujat puhuvat sekä omasta että tutkimuksen tekijän mielestä yleispuhekieltä, joskin yksi informanteista kokee, että hänen puheessaan kuuluu myös Oulun murre.

Tutkimuksen kannalta oleellista on yleispuhekielisyyden lisäksi se, että osallistujat toteuttavat spontaanipuheessaan rajapidennyksen e-loppuisten nominien yhteydessä – muutoin ei olisi perusteltua olettaa sitä epäsanojenkaan yhteydessä. Tämän vuoksi kaikki koehenkilöt olivat tutkimuksen tekijälle ennestään tuttuja. Tuttuudesta huolimatta osallistujat eivät olleet tietoisia tutkimuksen tarkoituksesta.

2.2 Koemateriaali

Pseudosanojen yhteydessä toteutuvaa tai toteutumatta jäävää rajapidennystä ei ole mahdollista tarkastella luonnollisesta spontaanipuheesta, joten tutkimusta varten laadittiin kuusisanaisia suomenkielisiä lauseita koehenkilöiden luettavaksi. Rajapidennyksen mahdollisesti laukaiseva herätesana sijaitsee lauseessa aina kolmantena, joten huomionkohteena on neljännen sanan alkukonsonantin piteneminen. Neljäs sana haluttiin vakioda mahdollisimman hyvin, joten se on aina kolmi- tai nelitavuinen, mutta monotonisuuden välttämiseksi sen äänteitä ei kohdeäännettä lukuun ottamatta rajoitettu. Koemateriaalia kehiteltäessä kiinnitettiin runsaasti huomiota siihen, ettei tutkimuksen tarkoitus paljastuisi koehenkilöille.

2.2.1 Herätesanat

Tutkimuksen lauseissa on neljänlaisia rajapidennyksen mahdollistavia herätesanoja, jotka eroavat toisistaan vain ensimmäisen äänteen osalta: **oikeita** e-loppuisia sanoja, joilla on tarkoitteensa reaalimaailmassa (*vene, kanne*), niitä mukailevia **epäsanoja** (*lene, sanne*), lainasanoja muistuttavia **epälainasanoja** (*grene, fanne*) ja **erisnimiä** (*Rene, Janne*). Taulukossa 2 on havainnollistettu herätesanojen keskinäistä vastaavuutta.

Epäsanalla, pseudosanalla tarkoitetaan tietyn kielen fono- ja morfotaksin mukaista sanaa, joka ei esiinny kielessä (Tieteen termipankki: *Nimitys:epäsana*). Se ei tarkoita mitään, mutta mukailee kielen sääntöjä sekä ortografiassa että ääntämisessä. Tässä työssä käytetään nimitystä **pseudosana** vain tarkoittamassa epäsanoina ja epälainasanoja **yhdessä**. Muuten pseudosanakategorian alaryhmistä puhutaan **epäsanoina** ja **epälainasanoina/epälainoina**.

Vokaali–konsonantti-rakenne sanoissa on kahdenmuotoinen: *CVCE ja *CVCCE, joissa

C = mikä tahansa konsonantti

CC = minkä tahansa konsonantin geminaatta

V = mikä tahansa vokaali

E = vokaali [e]

Asteriski * indikoi mahdollista toista konsonanttia, jolla tuotetaan lainasanamielleyhtymän aiheuttava konsonanttiklusteri sanan alkuun, kuten sanassa *grene*. Tavunalkuinen CC-klusteri ei ole suomelle tyypillinen, vaikka fonotaksi sen salliikin tietyissä, alun perin lainaa olevissa sanoissa, kuten *prinsessa* ja *kloonata* (mm. Verkkokielioppi). Sananalkuisten konsonanttiyhtymien lisäksi epälainasanoja ilmennetään käyttämällä niissä suomen fonologiseen järjestelmään kuulumattomia konsonantteja, kuten *z* ja *g*. Konsonanttiyhtymät ja vierasperäiset konsonantit mainitaan tyypillisinä lainasanan tuntomerkkeinä myös lähdekirjallisuudessa (Häkkinen, 1990, 261–263, viitattu lähteessä Koivisto, 2013, 202).

Erisnimi on ilmaus, jolla nimetään yksilö niin, että tämä voidaan tunnistaa muiden vastaavien joukosta (Kielitoimiston ohjepankki: *Erisnimi*). Tämän tutkimuksen nimikategoriaan sijoitettiin ”selvien” suomalaisten eris- ja lempinimien, kuten *Janne* ja *Late*, lisäksi ”epänimiä”, jotka voisivat olla – ja mahdollisesti ovatkin – etunimiä, sukunimiä tai lempinimiä. Tällaisia olivat esimerkiksi *Linne*, *Kuhe* ja *Rumme*, ja nämä kirjoitettiin isolla alkukirjaimella koelauseissa.

Taulukko 2. Esimerkkejä tutkimuksessa käytetyistä herätesanoista, jotka eroavat toisistaan ensimmäiseltä äänteeltään. *CVCE- ja *CVCCE -muotoiset sanat vuorottelevat taulukossa.

Oikea sana	Epäsana	Epälainasana	Nimi
vire	nire	skire	Mire
rinne	kinne	grinne	Linne
hake	lake	tsake	Make
murre	surre	zurre	Lurre
kate	sate	trate	Late
lumme	humme	grumme	Rumme

Neljää herätesanatyyppeä oli ensimmäisillä lukukerroilla 12 erilaista kutakin. Ne toistuivat lauseissa kolme kertaa, jolloin oikeita sanoja, nimiä, epäsanoja ja epälainoja oli lukukerralla 36. Ensimmäisten äänitysten jälkeen (ks. 2.2.4) oikeiden sanojen määrä tuplattiin: mukaan otettiin uusia oikeita e-loppuisia sanoja. Vertailu eri sanatyyppien välillä katsottiin kuitenkin hyvin mahdolliseksi, etenkin jos epäsanat ($n = 36$) ja epälainasanat ($n = 36$) lasketaan kuuluvaksi samaan pseudosanojen kategoriaan, jolloin saadaan 72 oikeaa sanaa ja 72 pseudosanaa lukukerralla.

2.2.2 Kohdeäänteet

Herätesanoja seuraavat sanat alkavat kolmella eri äänteellä, jolloin varsinainen huomionkohde, mahdollisesti rajapidentyvä konsonantti on [p], [m] tai [h]. Kohdeäänteiksi [p] ja [m] valikoituivat artikulaatiopaikkansa perusteella: molemmat ovat bilabiaalisia, jolloin voidaan olettaa, että niiden tuottamiseen osallistuvien artikulaattorien matka edeltävästä [e]-äänteestä ei juuri eroa, ja ne ovat tässä mielessä toisiinsa rinnastettavia. Äänteiden kestot ovat jossain määrin kielikohtaisia, mutta valtaosa tutkimuksista on osoittanut labiaaliset äänteet alveolaarisia ja velaarisia pidemmiksi, kun muut tekijät puheessa on vakioitu (Lehiste, 1970, 27). Soinnittomat äänteet, kuten klusiili [p], ovat puolestaan osoittautuneet useiden tutkimusten mukaan mitatuilta kestoiltaan pidemmiksi kuin soinnilliset, kuten nasaali [m].

[h] valikoitui kohdeäänteeksi erikoisuutensa vuoksi. Suomen [h]:lla ei ole selkeää ääntymäpaikkaa tai ilmenemismuotoa, ja se määritellään osan lähteistä mukaan soinnittomaksi frikatiiviksi, toisten mukaan se on useimmiten soinnillinen ja vain joissain tapauksissa soinniton. Soinnillisuuteen vaikuttaa äänneympäristö, [h] on soinnillinen esimerkiksi vokaalien välissä, kuten sanassa *piha*, soinniton puolestaan soinnittomien äänteiden yhteydessä, kuten sanassa *vihko*. Ominaisuuksiensa erityislaatuisuuden lisäksi [h]:n rajapitenemistä ei odoteta samalla tavalla kuin muiden äänteiden (mm. Ikola, 2001, 21).

Rajapidennysilmiötä tarkastellaan siis muuttumattomilta ominaisuuksiltaan toisistaan huomattavasti poikkeavien konsonanttien avulla. Foneettisten äännekestojen selvittämisen lisäksi tavoitteena on tarkastella ilmiötä myös kvalitatiivisesti: luokitella äänteet kategorioihin niiden havaitun keston mukaan ja vertailla luokkia toisiinsa.

2.2.3 Lauseet

Tutkimuksen analysoitavat lauseet ovat kuusisanaisia, herätesana sijaitsee lauseessa aina kolmantena ja kohdeäänteen sisältävä vähintään kolmitavuinen sana neljäntenä. Koelauseista tehtiin tarkoituksellisen pitkiä, jotta huomio lukiessa jakautuisi usealle sanalle, ja tutkimuskohde olisi näin varmemmin koehenkilön päättelemättömissä. Lauseet sisälsivät myös runsaasti nimiä, joiden kirjoitusasu oli lyhyt, mutta jotka oli mahdollista lausua lyhyenä tai pitkänä, kuten *Nora* ja *Lena*. Osa nimistä oli puolestaan kirjoitettu pitkään muotoon, kuten *Tiia* ja *Anniina*. Tällä haluttiin antaa kokeen tarkoitusta mahdollisesti pohtivalle koehenkilölle oletusvaihtoehto tutkimusaiheesta – ja samalla kiinnittää huomio toisaalle itse tutkittavasta aiheesta.

Koelauseiden rinnalle laadittiin täytelauseita, joiden tarkoitus oli tehdä lukukokemuksesta mielekkäämpi ja rikkoa lauseiden kaavamaisuutta. Täytelauseissa kiinnitettiin erityistä huomiota niiden foneettiseen tasapainoisuuteen; samassa lauseessa esiintyi rajattu määrä konsonantti–vokaali-vaihtelua, millä pyrittiin lisäämään lukemisen miellyttävyyttä.

Esimerkkejä täytelauseista ovat *Mies myönsi kirjoittaneensa kirjeet* ja *Harjoituksen sisältö on muuttunut alkuajoista*.

Kehyslauseita laadittiin aluksi 48 erilaista, joihin herätesanat sijoitettiin. Lauseet pyrittiin rakentamaan siten, että ne säilyivät sisältonsä puolesta mahdollisimman uskottavina, vaikka herätesanaa niissä vaihdettiin. Vaihtamisella haluttiin normalisoida lauseen merkityksen tai sen muiden äänteiden mahdollista vaikutusta ilmiöön. Herätesanojen paikkoja vaihdettiin siten, että tietyssä lauseessa esiintyi vain tietyn sanatyypin sanoja: oikeat sanat esiintyivät keskenään samoissa lauseissa ja epäsanat, epälainasanat ja nimet omissaan. Esimerkkejä saman sanatyypin sanojen vaihtelusta kehyslauseissa on taulukossa 3.

Taulukko 3. Esimerkkejä tutkimuslauseista. Numerot sulkeissa kertovat lukuhetkestä. Esimerkiksi *1: 1* kertoo kyseessä olevan ensimmäisen luentakierroksen ensimmäinen lause.

Herätesanojen vaihtelu kehyslauseissa
<p>Noralle suunnattu puhe perustui aiempiin tietoihin. (1: 1) Noralle suunnattu murre perustui aiempiin tietoihin. (2: 57) Noralle suunnattu kanne perustui aiempiin tietoihin. (3: 9)</p> <p>Sympaattisen hermoston sunne hupenee ajan kuluessa. (1: 13) Sympaattisen hermoston luhe hupenee ajan kuluessa. (2: 58) Sympaattisen hermoston sate hupenee ajan kuluessa. (3: 32)</p> <p>Saran mielestä fanne muistutti riikinkukon pyrstöä. (1: 23) Saran mielestä trate muistutti riikinkukon pyrstöä. (2: 54) Saran mielestä grumme muistutti riikinkukon pyrstöä. (3: 48)</p> <p>Äänestä hätääntynyt Nelle huitaisi kiivaasti kädellään. (1: 59) Äänestä hätääntynyt Jare huitaisi kiivaasti kädellään. (2: 16) Äänestä hätääntynyt Late huitaisi kiivaasti kädellään. (3: 47)</p>

Lauseiden semanttiseen sisältöön kiinnitettiin runsaasti huomiota. Etenkin oikeita *vene*-sanoja sisältävissä keskityttiin paljon siihen, että lauseella olisi jokin tarkoitus myös herätesanan

vaihduttua. Muuten vaarana saattaisi olla tahaton kategoriavaihto: lauseen tarkoituksen katoaminen saattaisi saada koehenkilön suhtautumaan oikeaankin sanaan epäsanana. Aihepiiri lauseissa oli vaihteleva ja sisältö osittain myös tarkoituksellisen epäjohdonmukaista.

Herätesana esiintyi yhden kerran kullakin lukukierroksella. Lauseet luettiin kolme kertaa, ja jokainen herätesana esiintyi kaikkiaan kolmessa eri kehyslauseessa siten, että se sai jälkeensä kerran kunkin kohdeäänteistä [p], [m] ja [h]. Tätä havainnollistetaan taulukossa 4. Lauseet esiintyivät eri järjestyksessä jokaisella lukukierroksella, mutta järjestys oli sama koehenkilöiden välillä.

Taulukko 4. Esimerkkejä tutkimuslauseista. Jokainen herätesana luettiin kolmesti, kerran kunkin kohdeäänteen yhteydessä eri kehyslauseessa.

Herätesanojen esiintymisympäristöjä
<p>Noralle suunnattu puhe perustui aiempiin tietoihin. (1: 1) Muistoista kumpuava puhe heijastuu myös arkeen. (2: 49) Baskia muistuttava puhe muodostuu kahdesta kielestä. (3: 51)</p> <p>Sympaattisen hermoston sunne hupenee ajan kuluessa. (1: 13) Lähes sortunut sunne purettiin viime viikolla. (2: 52) Varjoisten soiden sunne muistuttaa kukittuaan saniaista. (3: 37)</p> <p>Saran mielestä fanne muistutti riikinkukon pyrstöä. (1: 23) Kuohkeuden takaamiseksi fanne paistetaan vähän kerrallaan. (2: 33) Rakennuksen sisäkaton fanne helpottaa aikakauden tunnistamista. (3: 29)</p> <p>Äänestä hätääntynyt Nelle huitaisi kiivaasti kädellään. (1: 59) Tiian poistuttua Nelle muisteli kohtaamaansa ongelmaa. (2: 21) Ministerinä toimiessaan Nelle perusteli aina kantansa. (3: 6)</p>

Ensimmäisillä lukukerroilla yksi lukukierros sisälsi 60 lausetta siten, että neljän analysoitavan lauseen jälkeen seurasi aina täytelause. Kolmella lukukierroksella lauseita luettiin yhteensä 180, joista analysoitavia oli 144 ja täytelauseita 36. Täytelauseet olivat kaikki erilaisia keskenään, millä pyrittiin lisäämään vaihtelevuutta lukemiseen.

Ensimmäisten lukukertojen jälkeen analysoitavien lauseiden määrää lisättiin (ks. 2.2.4). Alkuperäisten rinnalle kehitettiin 36 uutta koelauseetta, ja täytelauseiden määrää vähennettiin. Ensimmäisen ja toisen kierroksen lausemääränä säilyi 60, kolmannella niitä oli 70. Lopuilla

lukukerroilla oli yhteensä 190 lausetta, joista 180 oli koelauseita. Koko tutkimuksessa analysoitavia lauseita oli yhteensä 2412, ja täytelauseineen kokonaismäärä oli 2630.

2.2.4 Pilottikoe, ensimmäiset äänitykset ja lukumateriaalimuutos

Lukumateriaalin toimivuutta haluttiin testata pilottikokeella ennen äänityksiä, joten osa lauseista esiluettiin kahdella lukijalla, joista toinen toimi myöhemmin myös tutkimuksen koehenkilönä. Huomiota kiinnitettiin erityisesti epäsanonja ja epälainasanoja sisältäviin lauseisiin ja siihen, kuinka koehenkilö suhtautui koemateriaaliin. Esilukijoiden pseudosana-lauseisiin väsymisen perusteella mielenkiintoa ylläpitäviä tarkoitussisältöisiä lauseita päätettiin lisätä, ja tämä tehtiin kasvattamalla täytelauseiden määrää alun perin aiotusta. Pilottikokeen perusteella oli muuten syytä olettaa, että koemateriaali toimisi tarkoitettussa tehtävässään.

Kun kolmea koehenkilöä oli äänitetty, ilmeni kuitenkin, että joka viidenneksi sijoitetuista täytelauseista huolimatta koehenkilöt vaikuttivat ”sokaistuvan” lukemalleen. Etenkin koetilanteen loppupuolella rajapidennys alkoi jäädä toteutumatta melko systemaattisesti myös sellaisissa oikeissa sanoissa, joissa se puhujien spontaanipuheessa toteutuu. Syyksi tähän arveltiin herätesanojen jakaumaa: koska sanatyypipedustus oli tasainen, oikeita *vene*-sanoja oli vain neljäsosassa koelauseita. Loput sisälsivät pseudosanoja ja nimiä, joiden yhteydessä rajapidennys ei olisi yhtä todennäköinen, ja tällä saattaisi olla vaikutus ilmiöön myös oikeissa sanoissa. Ja jos rajapidennys ei toteudu koehenkilölle tyypillisellä tavalla oikeissakaan sanoissa, voisiko sitä edes olettaa muiden sanatyypien yhteydessä?

Koemateriaalia muutettiin siten, että nimiä, epäsanonja ja epälainoja sisältävät lauseet säilytettiin sellaisinaan, mutta valtaosa täytelauseista korvattiin uusilla, oikeita *vene*-sanoja sisältävillä lauseilla. Samoja puhujia ei äänitetty enää uudestaan, joten ei ole varmaa, oliko koemateriaalimuutoksella vaikutusta. Muutoksen jälkeisissä äänityksissä ero kokeen alku- ja loppupuolen oikeissa sanoissa ei kuitenkaan ollut niin kuuluva, vaan rajapidennys reaalistui niissä enemmän informantin spontaanipuhetta vastaavalla tavalla.

2.3 Tutkimuskysymykset

Tämän tutkimuksen tavoite on selvittää, onko rajapidennys produktiivinen ilmiö. Sitä tarkastellaan erityisesti seuraavien tutkimuskysymysten osalta:

1. Onko rajapidennyksen toteutumisessa eroa tutkimuksen eri sanatyyppejen välillä?
2. Onko herätesanan rakenteella vaikutusta rajapidennyksen toteutumiseen?

Tutkimuskysymyksistä esitetään hypoteesit, joita tarkastellaan tilastollisin menetelmin:

1. H0: Rajapidennyksen toteutumisessa ei ole eroa tutkimuksen eri sanatyyppejen välillä.
H1: Rajapidennyksen toteutuminen on erilaista tutkimuksen eri sanatyyppejen välillä.
2. H0: Herätesanan rakenne ei vaikuta rajapidennyksen toteutumiseen.
H1: Herätesanan rakenne vaikuttaa rajapidennyksen toteutumiseen.

Tilastollisesti tarkasteltavien hypoteesien lisäksi esitetään myös tutkimuksen tekijän ennakkoletukset. Ne ovat ympäristöstä tehtyihin havaintoihin, äidinkielisen tutkijan intuitiiviseen kielikäsitukseen sekä jossain määrin aiempaan tutkimukseen (Lehtimäki, 2012) perustuvia.

E1: Rajapidennyksen toteutuminen on erilaista tutkimuksen eri herätesanatyyppejen yhteydessä, ja seuraavassa esitellään ennakkoletukset eri tyypeistä:

E1.1: Rajapidennys toteutuu tutkimuksen **oikeiden sanojen** yhteydessä. Kaikki oikeat sanat kuuluvat *vene*-ryhmään, jonka yhteydessä rajapidennyksen on mahdollista reaalistua.

E1.2: Rajapidennys toteutuu **epäsanan** yhteydessä, jos epäšana on tarpeeksi uskottava kontekstissaan. Jos se tuntuu keksityltä ja yhteyteen sopimattomalta, rajapidennys ei toteudu.

E1.3: Rajapidennys ei todennäköisesti toteudu **epälainasanojen** yhteydessä. Lainasanojen *psyky* ja *single* tavoin myös epälainojen oletetaan lukeutuvan *nalle*-ryhmään. Sanan rakenne saattaa kuitenkin vaikuttaa toteutumiseen (ks. E2).

E1.4: Rajapidennys ei toteudu **nimien** yhteydessä. Poikkeuksena saattaisivat olla nimet, joilla on myös jokin yleisnimimerkitys, kuten *Aarre*, mutta näitä tapauksia ei tässä työssä tutkita.

E2: Sanan rakenne vaikuttaa rajapidennyksen toteutumiseen pseudosanoissa, eli epä sanoissa ja epälainasanoissa. Yleisinhimillisen rytmillisyyden tavoittelun vuoksi rajapidennyksen oletetaan toteutuvan niissä useammin *CVCCE-muotoisten sanojen yhteydessä. Oikeiden sanojen ja nimien kohdalla eroa ei oleteta olevan.

Tutkimuskysymysten ja niistä muodostettujen hypoteesien ja ennakkoletusten lisäksi tutkimuksessa tarkastellaan kohdeäänteen vaikutusta rajapidennyksen toteutumiseen aineistolle tehdyn auditiivisen luokittelun avulla.

2.4 Äänitys

Äänitykset tehtiin Helsingin yliopiston äänitysstudioissa käyttäen kondensaattorimikrofonia (AKG C4000B). Mikrofoni sijoitettiin telineeseen kiinnitettynä hieman sivuun koehenkilön suusta, jotta mahdolliset klusiilien aiheuttamat ilmapurskahdukset, frikatiivit tai hengitys eivät aiheuttaisi häiriötä äänitykseen. Lisäksi käytettiin Universal Audio Apollo Twin MkII -A/D-muunninta ja Audacity-äänieditoria, jolla puhunnokset tallennettiin wav-muodossa tietokoneelle. Äänitykset tehtiin 16bit/48 kHz näytteenottotaajuudella.

Koehenkilö istui äänieristetyssä huoneessa ja luki lauseet edessään olevalle nuottitelineelle asetelluilta papereilta. Ohjeistuksena oli lukea lauseet käyttäen luontevalta tuntuvaa lukunopeutta, mutta säilyttäen kuitenkin selkeys puheessa. Informanttia ohjeistettiin, että osa lauseista sisältää vieraita sanoja tai epäsanvoja, mutta niitä ei tulisi paljastaa tauoilla, erityisillä painotuksilla tai poikkeavalla intonaatiolla. Koehenkilölle tähdennettiin, että tutkimuksen tavoite ei ole kartoittaa eläytymiskykyä tai saada luentaa kuulostamaan mielenkiintoiselta, joten puheen vaihtelevuudelle tai korostetulle prosodialle ei olisi tarvetta. Ohjeilla pyrittiin välttämään taukoja lauseessa, sillä tutkimuskohteena on yhtäjaksoisen puheen ilmiö.

Koehenkilöä ohjeistettiin lukemaan lause kokonaan uudestaan, jos hän teki lukiessaan virheen tai takelteli sanoissaan. Tarvittaessa informantille oli mahdollista puhua tarkkailuhuoneesta ja antaa lisäohjeita tai pyytää uudelleen lukemista. Lauseet äänitettiin kolmessa erässä, ja äänityskertojen välissä pidetyn tauon aikana tarvittaessa tarkennettiin ohjeita.

Lauseita lukiessaan moni puhuja piti, luultavasti huomaamattaan, hyvin, hyvin lyhyen tauon juuri huomionkohteena olevalla sananrajalla. Tätä tapahtui erityisesti silloin, kun herätesanana oli pseudosana, mutta myös muulloin. Tauko aiheuttaisi vääjäämättä sen, ettei lausetta voisi analysoida, joten äänityskierrosten välissä informanteille korostettiin olevan tärkeää, ettei minkäänlaisia taukoja luennassa esiintyisi. Kokeen tarkoituksen paljastumattomuuden vuoksi ei kuitenkaan voitu painottaa tämän tärkeyttä juuri kolmannen ja neljännen sanan välillä. Liiallinen tauottomuuden vaatiminen olisi sitä paitsi saattanut johtaa ”konemaiseen” lukemiseen, sanojen latelemiseen vailla yritystä ymmärtää lukemaansa. Tämän lisäksi tauot olivat niin huomaamattomia, ja luultavasti lukijan itsensä tiedostamattomia, ettei niitä eitutkimuksellisessa mielessä edes mieltäisi tauoiksi. Niinpä toistuva huomauttelu niistä olisi luultavasti aiheuttanut lukijalle epämukavuutta – poistamatta taukoja kuitenkaan.

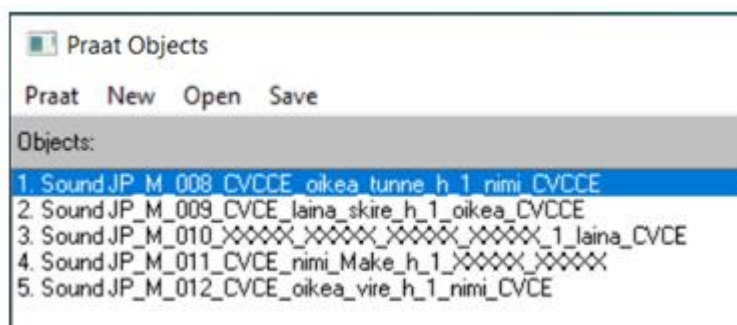
2.5 Aineiston käsittely

Tutkimusaineisto koostui 2630 lauseesta, joista analysoitavia oli 2412. Lauseiden muuttujina olivat sanatyypit (neljä erilaista: oikea, epäsanana, epälainasana, nimi), sanan rakenne (kaksi erilaista: *CVCE, *CVCCE) ja kohdeäänne (kolme erilaista: [p, m, h]). Oikeita sanoja oli aineistossa ensimmäisten äänitysten jälkeisen muutoksen seurauksena 900, muita tyyppejä 504 kutakin. Epäsanat ja -lainat on mahdollista lukea samaan pseudosanojen ryhmään, jolloin niitä oli yhteensä 1008, mikä on hyvin verrattavissa oikeisiin sanoihin. *CVCE- ja *CVCCE -muotoisia sanoja oli tutkimuksessa 1206 kumpiakin. Myös kohdesegmenttien jakauma aineistossa oli tasainen: kutakin äännettä oli 804.

Hylättyjä lauseita oli aineistossa lopulta 540 (ks. 2.5.2). Näistä 155 sisälsi oikean sanan, 110 nimen, 126 epäsanana ja 149 epälainasanan. Analysoitavaksi jäi siis lopulta 745 oikeaa sanaa, 733 pseudosanaa (378 epäsanana ja 355 epälainasanaa) ja 394 nimeä: yhteensä 1872 lausetta herätesanoineen. Lukumateriaalimuutosta edeltäviä lauseita, joissa rajapidennys toteutui selvästi puhujan spontaanipuheesta eroavalla tavalla, ei hylätty.

2.5.1 Akustinen analyysi

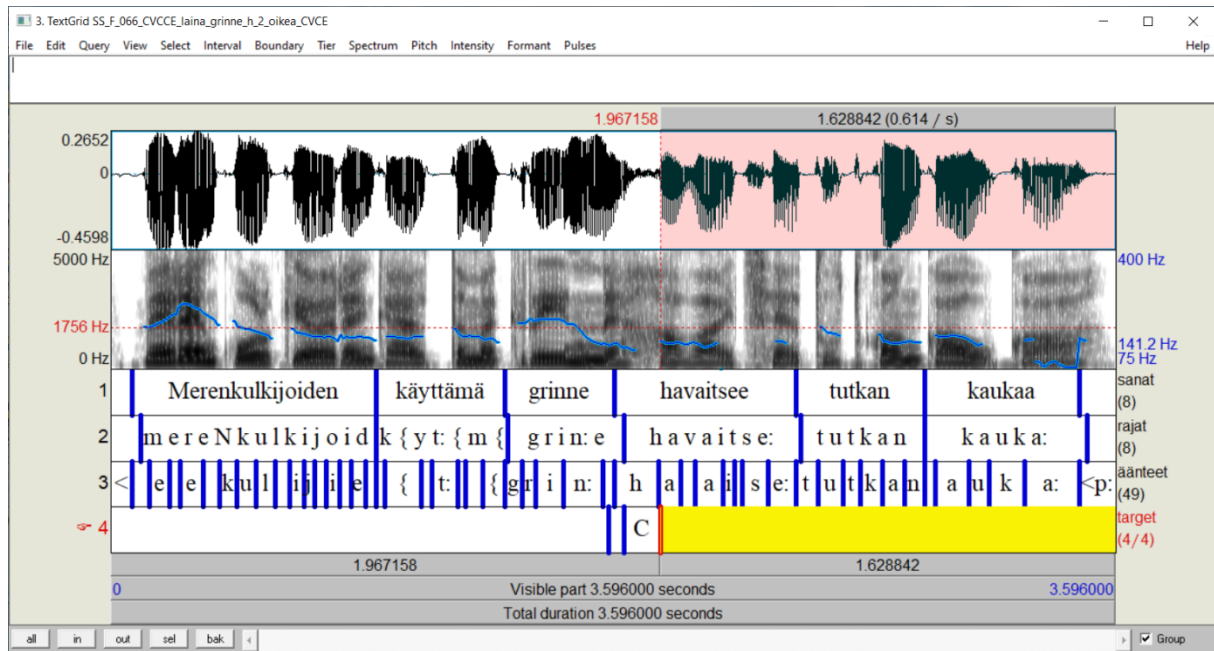
Tutkimuksen aineisto käsiteltiin Praatilla (Boersma & Weenink, 2019). Koehenkilöiltä äänitetyt puhunnokset olivat raakamuodossaan kolmessa eri wav-tiedostossa, joista jokainen sisälsi yhden kolmesta äänityskerrasta, 60 tai 70 lausetta. Äänitiedostoista poistettiin virheelliset luennat, yskähdykset ja mahdolliset muut häiriöäänet, ja lauseiden väliin merkittiin tauot. Puhujien tuottamat lauseet pilkottiin yhtenäisistä äänitiedostoista yksittäisiksi wav-tiedostoiksi, ja ne nimettiin informatiivisesti: nimi sisälsi tiedot puhujasta, herätesanasta ja kohdeäänteestä. Nimitietoihin lisättiin myös kyseistä lausetta edeltävän lauseen herätesanan tyyppi ja muoto, sillä koetta suunniteltaessa edellisellä lauseella arveltiin saattavan olla *priming*-vaikutus seuraavaan, mikä haluttaisiin mahdollisesti huomioida analyysissä. *Priming*-ilmiössä edeltävä ärsyke vaikuttaa seuraavan ärsykkeen käsittelyyn, ja edeltävän lauseen herätesanalla otaksuttiin mahdollisesti olevan vaikutus erityisesti pseudosanoja luettaessa. Kuvassa 1 on havainnollistettu lauseiden nimeämistä.



Kuva 1. Koehenkilön JP viisi tutkimuslausetta Praatin objekti-ikkunassa. Valitun tiedoston nimestä selviää, että puhuja on mies (M), lause on järjestyksessä kahdeksas hänen lukemansa (008), herätesana on muotoa *CVCCE ja tyyppiä **oikea**, sana on *tunne*, kohdeäänne on [h], puhuja lukee sanan ensimmäistä kertaa (1), ja sitä edeltää sana tyyppiä **nimi** muodoltaan *CVCCE. XXXXX-merkitty lause nro10 on täytelause. Lauseen 11 nimessä edeltävä sana on siis myös XXXXX-merkitty, sillä täytelauseilla ei oleteta olevan *priming*-vaikutusta.

Annotoinnin nopeuttamiseksi apuna käytettiin pakotettuun kohdistukseen perustuvaa WebMAUS Basic -palvelua (Schiel, 1999; Kisler, Reichel & Schiel, 2017). Äänitiedostot syötettiin ohjelmaan niitä vastaavien tekstitiedostojen kanssa, ja ohjelma loi kolmikerroksisen TextGrid-tiedoston kustakin wav-tiedostosta kohdentaen grafeemit niitä vastaaviin äänteisiin. Tarkkuutta vaativassa äännekestojen tutkimuksessa on välttämätöntä tarkastaa segmentointi myös käsin, joten jokainen TextGrid käytiin läpi manuaalisesti. Lauseen alku, loppu, sekä kolmannen ja neljännen sanan rajat korjattiin oikeille kohdilleen, ja pitkistä klusiileista aiheutuneet tauot poistettiin. Lisäksi luotiin neljäs kerros, jolle rajattiin tarkasti herätesanan [e] sekä kohdeäänne, eli neljännen sanan alkukonsonantti (kuva 2).

Koska 2412 puhunnoksen manuaalinen annotointi on verrattain hidasta yhden henkilön tekemänä, haluttiin varmistua, että ensimmäisinä segmentoidut TextGridit vastaisivat viimeisimpinä tehtyjä. Segmentoidessa oli määritelty tietyt reunaehdot äännerajojen asettamiselle, mutta ensimmäisten ja viimeisimpien annotointien välillä kului useita kuukausia. Systemaattisuuteen pyrkimisestä huolimatta ensimmäisinä ja viimeisimpinä annotoitujen TextGridien välillä oli jonkin verran eroavaisuuksia, joten kaikki tiedostot käytiin uudestaan läpi, ja niihin tehtiin tarvittavat korjaukset.



Kuva 2. Praatin analyysi-ikkunassa naispuhujaa SS:n lukema lause herätesananaan epälaina *grinne*. WebMAUS on luonut kolme ylintä kerrosta, joista toiselle sanojen rajat on korjattu käsin oikeille paikoilleen. Neljäs, kohdeäännekerros on luotu manuaalisesti, ja sille on tarkasti rajattu herätesanan [e] (merkinnällä V) ja kohdeäänne [h] (merkinnällä C). Kuvasta ilmenee myös, että rajapidennys [h]:ssa on toteutunut, sillä frikatiivihäly on nähtävissä.

TextGrideihin annotoitujen äännekestoja keräämiseen käytettiin Python-skriptiä, joka laski kestot neljännen kerroksen [e]:lle ja kohdekonsonantille sekä jokaiselle lauseen kuudesta sanasta. Nämä taulukoitiin Exceliin. Tämän lisäksi taulukkoon merkittiin käsin jokaisesta lauseesta tehty kuulonvarainen pituusluokittelu (ks. 2.5.2.). Omaan sarakkeeseensa kirjattiin myös, oliko lukiessa tapahtunut virhe.

2.5.2 Auditiivinen analyysi

Praatilla tehdyn akustisen analyysin lisäksi aineistolle tehtiin myös kuulonvaraisuuteen perustuva auditiivinen luokittelu kohdeäänteen pitenemisestä (ks. myös Karlsson & Lehtonen, 1977, 9–10). Tutkimuksen tekijä kuunteli aineiston läpi kaikkiaan neljä kertaa. Aikomus oli luokitella piteneminen kahteen luokkaan sen mukaan, toteutuiko rajapidennys vai ei, mutta kuunnellessa ilmeni, että jako ei kaikkien puhunnosten kohdalla ollut yksiselitteinen. Osassa puhunnoksia kohdeäänne ei toteutunut pitkänä, eikä ilmiö jäänyt täysin toteutumatta, vaan havaittavissa oli jotakin siltä väliltä. Niinpä päädyttiin kolmiluokitteluun. Ensimmäinen kuuntelu tehtiin ennen Praat-analyysiä, jolloin paperille merkittiin jokaisen kohdeäänteen

havaittu pituus: lyhyt (merkintä Q1), puolipitkä (Q2) tai pitkä (Q3). Akustisen analyysin ja segmentoinnin tarkistuskierron yhteydessä pituuksia arvioitiin uudestaan. Takeltelut, tauot ja empimiset kirjattiin yhteisellä virhemerkinnällä paperille ja Praatin kohdeäännekerrokselle.

Annotoinnin valmistuttua tehtiin viimeinen auditiivinen arviointi. Sillä pyrittiin sulkemaan pois sekä Praatin analyysi-ikkunan tuoma visuaalinen informaatio että tutkimuksen tekijän mahdollinen asenteellinen suhtautuminen; olihan mahdollista, että kuulijan olettamus ilmiön toteutumisesta vaikuttaisi myös hänen havaintoonsa siitä. Tämä saattaisi ilmetä erityisesti oikeiden sanojen kohdalla. Viimeinen arviointi tehtiin kuuntelemalla kunkin koehenkilön koko äänityskierros kerrallaan ja tekemällä merkinnät puhtaalle paperille. Kohdeäännettä pyrittiin kuuntelemaan mahdollisimman objektiivisesti keskittyen vain sen pituuteen. Samalla kiinnitettiin huomiota siihen, vaikuttivatko herätesanat toistensa ääntämiseen.

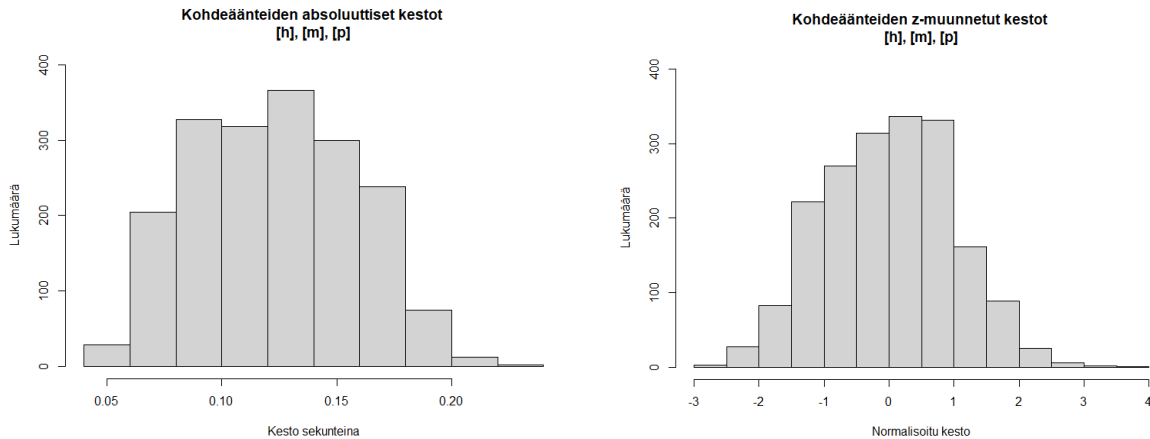
Äänityksiä kuunnellessa ja akustisesti analysoitaessa huomattiin, että useissa äänityshetkellä sujuvilta vaikuttaneissa puhunnoksissa oli hyvin pieni tauko juuri huomionkohteena olevalla sananrajalla. Joissain tapauksissa se johtui epäröinnistä, toisinaan lukija puolestaan rytmitti puhunnoksen lukien sen ikään kuin kahtena erillisenä lauseena. Tämän kuuli painotuksesta: lausepaino lankesi sekä lauseen ensimmäiselle että neljännelle sanalle, mikä synnytti kolmannen ja neljännen sanan rajalle tauon. Vaikka herätesanoilla ei havaittu olevan *priming*-vaikutusta toisiinsa, tällä ”uutistenlukijanuotilla” tuntui sellainen olevan; kun puhuja alkoi toteuttaa tätä puhetapaa, se jatkui pitkän aikaa.

Myös nämä erittäin lyhyitä taukoja sisältävät lauseet saivat virhemerkinnän. Virheellä merkittyjä oli analysoitavasta 2412 lauseesta lopulta 540, ja ne hylättiin. Vaikka hylkääminen tuntui joissain tapauksissa hieman liioitellulta – sillä puhunnoksesta oli useimmiten kuultavissa, tarkoittiko puhuja pidentää kohdeäännettä vai ei – siihen päädyttiin, sillä niiden analysoiminen olisi aiheuttanut virheellisiä kestoja.

2.6 Tilastolliset testit

Tilastolliset testit ovat apukeinoja, joilla on mahdollista testata tutkittavien hypoteesien paikkansapitävyyttä (Nummenmaa, 2004, 145). Tilastollista tarkastelua varten Excel-taulukko siirrettiin Rstudioon (Rstudio Team, 2020), joka on R-ohjelmointikieleen (R Core Team, 2020) pohjautuva tilastollisen laskennan käyttöympäristö. Jotta saataisiin selville, voiko tilastollisissa analyysissä käyttää tehokasta, mutta aineiston normaalijakaumaa edellyttävää parametrista testausmenetelmää, aineiston jakaumaa tutkittiin ensin kuvaajin. Kuvan 3

histogrammin (vasemmanpuoleinen) perusteella normaalius olisi mahdollinen, kvantiili–kvantiili-kuvaajan mukaan aineisto ei näyttänyt normaalisti jakautuneelta, joten oli syytä tehdä myös normaalisuustesti. Aineisto testattiin sekä Shapiro-Wilkin että suurille aineistoille ($n > 50$) suositellulla (Nummenmaa, 2004, 144) Kolmogorov-Smirnovin normaalisuustestillä, mutta kummankaan mukaan kyseessä ei ollut normaalisti jakautunut aineisto ($p < 0,05$).



Kuva 3. Kaksi histogrammia aineiston kohdeäänteistä. Vasemmanpuoleisessa konsonanttien absoluuttiset kestot: vaaka-akselilla kesto sekunteina, pystyakselilla äänteiden lukumäärä. Oikealla z-muunnetyt kestot: vaaka-akselilla normalisoitu kesto, pystyakselilla äänteiden lukumäärä.

Alun perin suunnitellun, kohdeäänteiden lause- ja sanakestoihin suhteuttamisen sijaan aineistolle tehtiin z-muunnos, jolla häivytettiin koehenkilöiden puh nopeuksista johtuvia eroja äänteiden pituuksissa. Muunnoksessa kohdeäänteiden puhujakohtaisista absoluuttisista kestoista vähennettiin puhujakohtainen keston keskiarvo, minkä jälkeinen erotus jaettiin absoluuttisten kestojen puhujakohtaisella keskihajonnalla. Z-muunnoksella normaalijakautuneesta muuttujasta tulee standardinormaalijakautunut, mutta se ei muuta ei-normaalisti jakautunutta muuttujaa normaalisti jakautuneeksi. Myös standardoitu aineisto testattiin kuvaajin sekä normaalisuustestein. Histogrammi (kuva 3, oikeanpuoleinen) sekä Kolmogorov-Smirnovin testi viittasivat siihen, ettei aineisto normalisoitunakaan täysin noudattanut normaalijakaumaa, vaikka merkitsevyystaso ($p = 0,0267$) olikin suhteellisen lähellä raja-arvoa 0,05.

Aineiston ollessa riittävän suuri, vaatimuksen normaalisti jakautuneisuudesta ei tarvitse täyttyä, jotta parametrisia testausmenetelmiä voidaan käyttää (mm. Gries, 2013, 33).

Toisaalta, aineiston testaaminen tehokkaalla, mutta mahdollisesti tulosta vääristävällä menetelmällä ei ole tarkoituksenmukaista (mts. 46). Koska tässä tutkimuksessa aineisto on suhteellisen suuri, mutta ei vaikuta täysin normaalisti jakautuneelta, sitä päädyttiin lopulta tarkastelemaan sekä parametrisin että epäparametrisin menetelmin – mikä on myös eräs suosituksista (mts. 160). Epäparametrisia menetelmiä voidaan soveltaa useammassa tilanteessa kuin parametrisia, mutta ne eivät ole yhtä tehokkaita hylkäämään nollahypoteesia silloin, kun se ei pidä paikkaansa (Nummenmaa, 2004, 268).

Lineaarinen regressioanalyysi on menetelmä, jolla voi tarkastella yhden tai useamman riippumattoman muuttujan vaikutusta riippuvaan muuttujaan. Normaalijakautuneisuuden lisäksi edellytyksenä on, että selitettävä muuttuja on vähintään välimatka-asteikollinen. (Nummenmaa, 2004, 317.) Menetelmällä voidaan tutkia yhtäaikaaisesti useamman tekijän vaikutusta ja selvittää yksittäisen muuttujan vaikutusosuus, kun muut vaikuttavat tekijät on huomioitu. Toinen tapa tarkastella tätä aineistoa on varianssianalyysi, joka sekin edellyttää normaalijakautuneisuutta ja vähintään välimatka-asteikollisuutta, ja myös sillä voi arvioida useamman tekijän vaikutusta riippuvaan muuttujaan. Välimatka-asteikolla mitataan numeerisia muuttujia, joiden etäisyydet toisistaan on mahdollista ilmoittaa. Tämän tutkimuksen aineiston selitettävä muuttuja on suhdeasteikollinen, eli se saa numeerisia arvoja, ja asteikon nollakohta on motivoitu. Välimatka-asteikolliselle aineistolle sopivat menetelmät sopivat myös suhdeasteikolliselle. (mts. 34–37.)

Koska tutkimuksessa haluttiin käyttää myös epäparametrisia menetelmää, näiden lisäksi päädyttiin Kruskal-Wallis testiin. Sen lisänä käytettiin Wilcoxonin testiä, toistettujen mittauksen t-testin epäparametrisia vastinetta, jolla on mahdollista selvittää, minkä muuttujaparien välillä eroa on. Monivertailujen yhteydessä käytetään p-arvojen korjausmenettelyä, jolla voidaan säilyttää alkuperäinen tilastollinen merkitsevyystaso monivertailuista huolimatta. Wilcoxonin testin kanssa käytettiin Bonferroni-korjausta, jolle lasketaan merkitsevyystaso jakamalla tavoiteltu merkitsevyystaso testien lukumäärällä, ja näin saatua arvoa käytetään jokaisessa yksittäisessä testissä. (Gries, 2013, 273.) Muuten tilastollisten testien yhteydessä käytettiin seuraavia merkitsevyysarvoja:

$p < 0,001^{***}$ = tilastollisesti erittäin merkitsevä

$p < 0,01^{**}$ = tilastollisesti merkitsevä

$p < 0,05^{*}$ = tilastollisesti melkein merkitsevä

$p < 0,1$ = tilastollisesti suuntaa antava

3 Tulokset

Tässä tutkimuksessa tarkastelun kohteena on rajapidennyksen toteutuminen ja ilmiön produktiivisuus sekä eri muuttujien vaikutus siihen. Tutkimuskysymyksistä muodostettuja hypoteeseja testattiin tilastollisin menetelmin. Menetelminä käytettiin sekä parametrisia että epäparametrisia testejä, jotta saatu tulos olisi aineiston jakautuneisuudesta huolimatta varmasti luotettava. Analyysissä käytettiin sekä normalisoitua (z-muunnettua) että normalisoimatonta aineistoa. Tarkastelukohteina olivat sanatyypin, sanan rakenteen ja kohdeäänteen vaikutus tutkittavaan ilmiöön, joista sanatyypin vaikutusta pidettiin produktiivisuuden kannalta keskeisimpänä. Myös auditiivisen arvioinnin perusteella muodostettuja kestokategorioita vertailtiin toisiinsa. Vertailtavina olivat kategorioiden kestot sanatyypeittäin ja äänteiden sijoittuminen kategorioihin eri sanaryhmien yhteydessä.

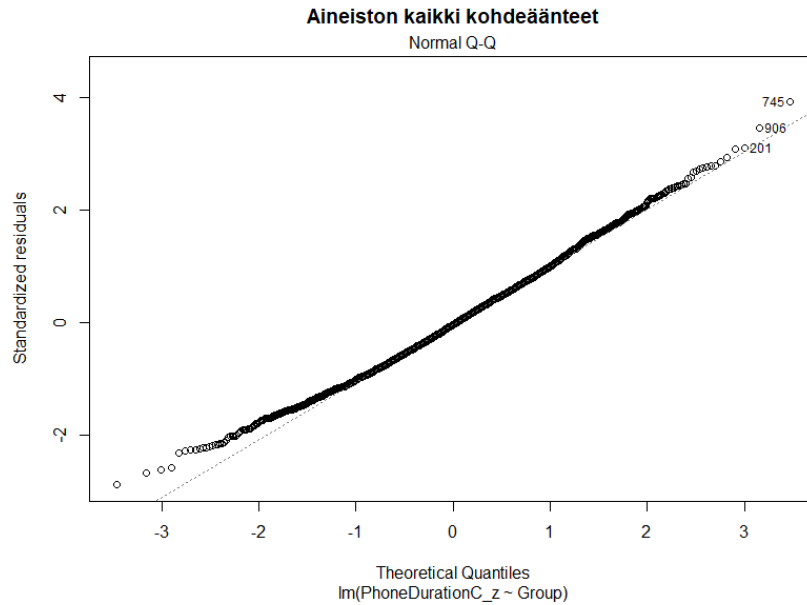
3.1 Sanatyypin vaikutus rajapidennyksen toteutumiseen

Herätesanatyypin vaikutusta ilmiön toteutumiseen tarkasteltiin kaikkiaan neljällä tapaa. Menetelminä käytettiin parametrissa lineaarisesta regressioanalyysistä ja epäparametrissa Kruskal-Wallis testin ja Wilcoxonin testin kera, ja aineistoa analysoitiin z-muunnettuna ja normalisoimattomana.

3.1.1 Lineaarinen regressioanalyysi: z-muunnettu aineisto

Sanatyypin välistä eroa tutkittiin lineaarisella regressioanalyysillä, jossa huomioon otettiin sanatyypin lisäksi kohdeäänteen. Ensin selvitettiin, onko oikeiden sanojen ja pseudosanojen (epäsanat ja epälainasanat yhdessä) välillä eroa rajapidennyksen aiheuttamisessa. Mallin mukaan pseudosanan jälkeinen konsonantti oli tilastollisesti erittäin merkittävästi oikean sanan jälkeistä lyhyempi ($\beta = -0,2796$, $t(2,1863) = -3,423$, $p = 0,0006^{***}$). Lineaarisen regression selitysaste oli tilastollisesti erittäin merkittävä ($R^2 = 0,16$, $F(8,1863) = 42,6$, $p < 0,0001^{***}$).

Pseudosanojen ryhmä koostui sekä kotoperäisistä äänteistä muodostuvista epäsanoina että vierasperäisistä äänteistä ja sananalkuisia konsonanttiyhtymiä sisältävistä epälainasanoista, joten näiden sanatyypin eroa oikeisiin sanoihin haluttiin testata erikseen. Eroa rajapidennyksen toteutumisessa vertailtiin kaikkien tutkimuksen sanatyypin kesken: oikeat sanat, epäsanat, epälainasanat ja nimet. Kuvassa 4 näkyy z-muunnetun aineiston kaikkien 1872 kohdeäänteen ([h, m, p]) kestojen asettuminen aineistoon sovitetulle suoralle.



Kuva 4. Kvantiili–kvantiili-kuvaaja z-muunnetusta aineistosta. Kohdeäänteiden voi nähdä sijoittuvan melko lähelle aineistoon sovitettua suoraa. Kuvassa kaikki kohdeäänteet ($n = 1872$).

Lineaarisen regressioanalyysin perusteella epälainasanat aiheuttivat oikeisiin sanoihin verrattuna tilastollisesti erittäin merkitsevästi lyhyemmän rajapidennyksen ($\beta = -0,5268$, $t(3,1860) = -5,353$, $p < 0,0001^{***}$). Myös nimet vaikuttivat rajapidennykseen oikeisiin sanoihin nähden tilastollisesti erittäin merkitsevästi lyhentävästi ($\beta = -0,7953$, $t(3,1860) = -8,234$, $p < 0,0001^{***}$). Epäsanat puolestaan erosivat oikeista sanoista vain vähän. Niiden aiheuttama rajapidennys oli oikeisiin sanoihin nähden hieman lyhentävä, mutta ei tilastollisesti merkitsevä ($\beta = -0,0344$, $t(3,1860) = -0,351$, $p = 0,7259$).

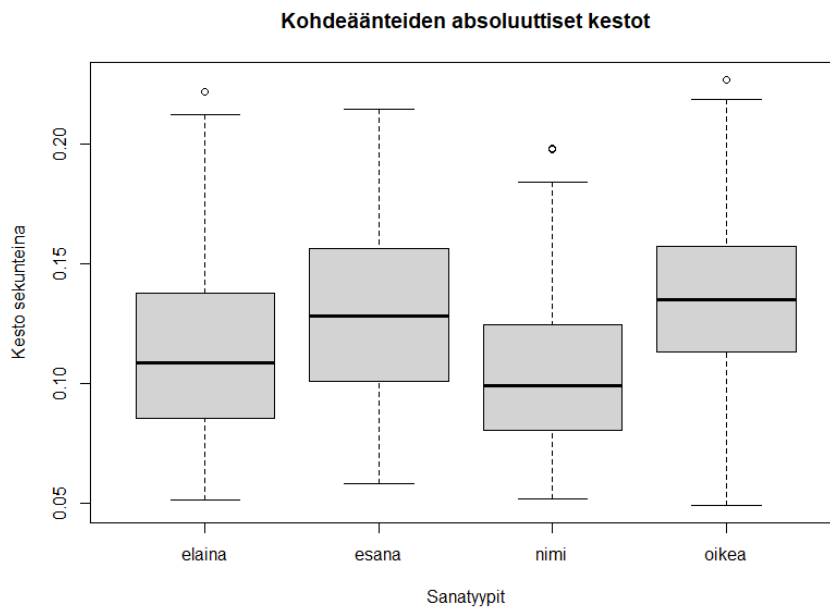
Sanatyypejä verrattiin toisiinsa myös asettaen vertailusanatyyppiksi nimet. Epäsanat erosivat nimistä aiheuttaen niitä tilastollisesti erittäin merkitsevästi pidemmän rajapidennyksen ($\beta = 0,7609$, $t(3,1860) = 6,763$, $p < 0,0001^{***}$). Epälainasanojen ero nimiin nähden oli hieman pidentävä ja tilastollisesti melkein merkitsevä ($\beta = 0,2686$, $t(3,1860) = 2,382$, $p = 0,0173^*$). Myös epäsanaja ja epälainasanoja vertailtiin toisiinsa. Epäsanat aiheuttivat epälainasanoja pidemmän rajapidennyksen, ja sanatyyppien välinen ero osoittautui tilastollisesti erittäin merkitseväksi ($\beta = 0,4924$, $t(3,1860) = 4,316$, $p < 0,0001^{***}$).

Lineaarisen regressiomallin selitysaste z-muunnetulle aineistolle oli tilastollisesti erittäin merkitsevä ($R^2 = 0,19$, $F(11,1860) = 39,6$, $p < 0,0001^{***}$).

3.1.2 Lineaarinen regressioanalyysi: absoluuttiset kestot

Lineaarisella regressiolla analysoitiin myös normalisoimaton aineisto. Oikeiden sanojen aiheuttama rajapidennys oli epälainasanoihin nähden tilastollisesti erittäin merkitsevästi pidempi ($\beta = 0,0169$, $t(3,1860) = 5,866$, $p < 0,0001^{***}$). Samoin oli oikeiden sanojen aiheuttama kesto nimien aiheuttamaan nähden ($\beta = 0,0264$, $t(3,1860) = 9,334$, $p < 0,0001^{***}$). Oikeiden sanojen ja epäsanojen välillä ei ollut absoluuttiskestoisinakaan tilastollisesti merkitsevää eroa ($\beta = 0,0017$, $t(3,1860) = 0,605$, $p = 0,5451$). Aiempaa mukaillen epäsanat erosivat tilastollisesti erittäin merkitsevästi nimistä ($\beta = 0,0247$, $t(3,1860) = 7,485$, $p < 0,0001^{***}$). Epälainojen ja nimien suhde erosi hieman normalisoidusta aineistosta ollen tilastollisesti merkitsevä ($\beta = 0,0095$, $t(3,1860) = 2,876$, $p = 0,0041^{**}$). Epäsanoissa rajapidennys oli epälainoja tilastollisesti erittäin merkitsevästi pidempi ($\beta = 0,0151$, $t(3,1860) = 4,540$, $p < 0,0001^{***}$). Kuvassa 5 näkyvät kohdekonsonanttien ([h, m, p]) kestot eri sanatyypin yhteydessä.

Lineaarisen regressiomallin selitysaste absoluuttisten kestojen aineistolle oli tilastollisesti erittäin merkitsevä ($R^2 = 0,43$, $F(11,1860) = 128,3$, $p < 0,0001^{***}$).

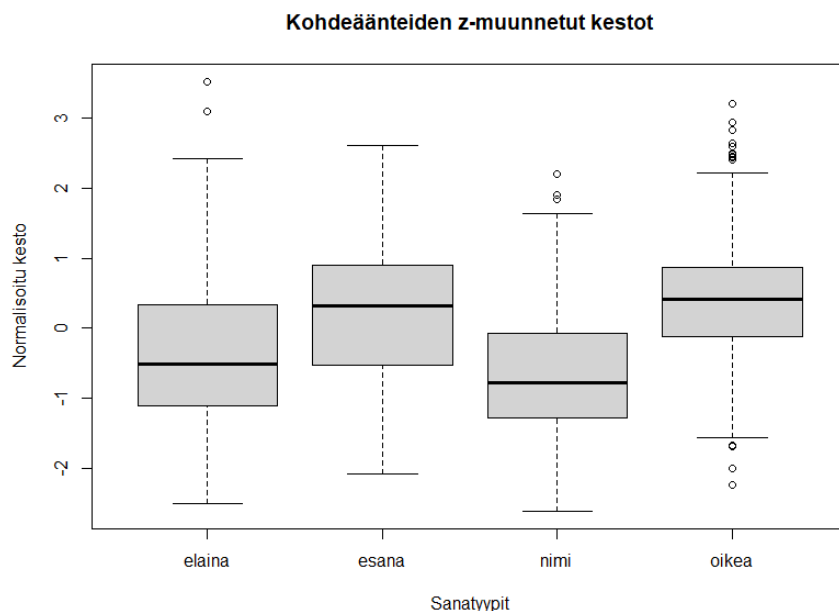


Kuva 5. Laatikko–jana-kuvaajat kaikkien kohdekonsonanttien ($n = 1872$) absoluuttisista kestoista sanatyypeittäin. Sanatyypit vasemmalta oikealle: epälainasana, epäsana, nimi, oikea sana. Laatikon keskellä oleva musta viiva on keston mediaani, laatikon alareuna on alakvartiiliin (25%) reuna, yläreuna yläkvartiiliin (75%) reuna: laatikoissa on puolet aineiston havainnoista. Kvartiileista piirretyt janat kertovat minimin ja maksimin. Aineiston muista arvoista huomattavasti poikkeavat arvot on merkitty janojen ulkopuolelle pienillä ympyröillä.

3.1.3 Kruskal-Wallis test: z-muunnettu aineisto

Aineistoa analysoitiin myös jakaumaoletuksettomalla testillä. Koska parametrisena menetelmänä olisi voinut käyttää myös varianssianalyysiä, epäparametrisena testinä käytettiin sen vastinetta, Kruskal-Wallis testia. Sanatyypien välisen eron osoittauduttua tilastollisesti erittäin merkitseväksi tehtiin Wilcoxonin testi, jotta saataisiin selville, minkä muuttujaparien välillä eroa on. Sen kanssa käytettiin Bonferroni-korjausta, joka alentaa merkitsevyysrajaa sen mukaan, kuinka monta monivertailua tehdään. Tässä monivertailussa tehtiin kuusi vertailua, joten merkitsevyysrajaksi saatiin 0,008.

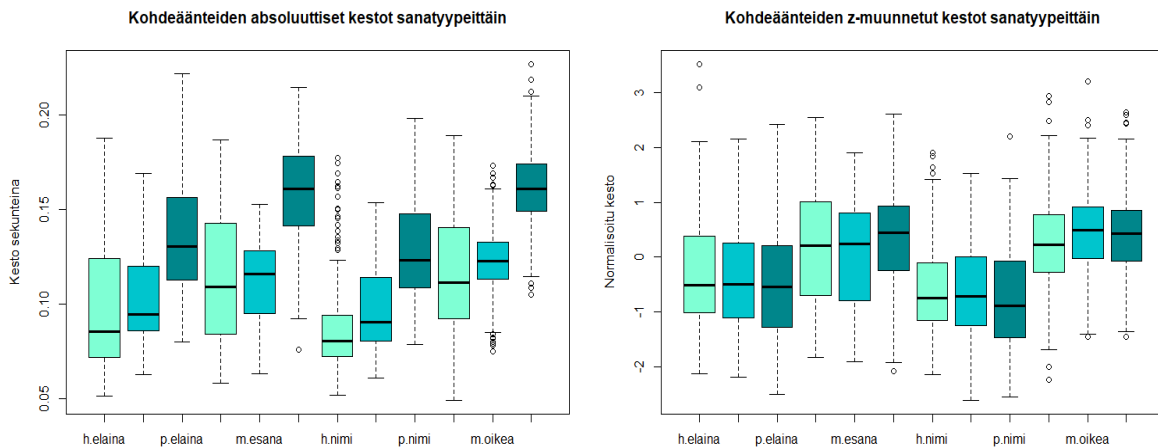
Tulos oli lineaarisen regressioanalyysin kanssa yhtenevä: oikeat sanat aiheuttivat nimiä ja epälainasanoja pidemmän rajapidennyksen, ja erot olivat tilastollisesti erittäin merkitseviä ($p < 0,0001^{***}$). Oikeiden sanojen ja epäsanojen välillä ei ollut tilastollisesti merkitsevää eroa p-arvon ollessa 0,0887. Nimet ja epäsanat erosivat toisistaan tilastollisesti erittäin merkitsevästi ($p < 0,0001^{***}$) ja nimet ja epälainas sanat toisistaan tilastollisesti merkitsevästi ($p = 0,0026^{**}$). Epäsanojen ja epälainasanojen ero oli tilastollisesti erittäin merkitsevä ($p < 0,0001^{***}$). Kuva 6 havainnollistaa rajapidennyksen toteutumista eri sanatyypien yhteydessä normalisoidussa aineistossa.



Kuva 6. Kuvassa kaikkien kohdekonsonanttien ($n = 1872$) z-muunnetyt kestot sanatyypeittäin. Sanatyypit vasemmalta oikealle: epälainasana, epäšana, nimi, oikea sana. Oikeiden sanojen kohdeäännekestojen voi nähdä jakautuvan hieman tiiviimmälle alueelle muihin nähden ja kestojen mediaanin sijoittuvan keskemälle laatikkoa kuin muissa sanatyypeissä.

3.1.4 Kruskal-Wallis test: absoluuttiset kestot

Myös normalisoimaton aineisto analysoitiin Kruskal-Wallis ja Wilcoxonin testein Bonferroni-korjauksella, ja tulokset olivat yhteneviä aikaisemmin saatujen kanssa. Oikeiden sanojen aiheuttama rajapidennys oli nimien ja epälainojen aiheuttamiin nähden tilastollisesti erittäin merkitsevästi pidempi ($p < 0,0001^{***}$). Oikeilla sanoilla ja epäsanoina ei ollut tilastollisesti merkitsevää eroa p-arvon ollessa 0,1747. Aiempia tuloksia mukaillen nimet ja epäsanat erosivat toisistaan tilastollisesti erittäin merkitsevästi ($p < 0,0001^{***}$) ja nimet epälainasanoista tilastollisesti merkitsevästi ($p = 0,0042^{**}$). Epäsanojen ja epälainasanojen välinen ero oli tilastollisesti erittäin merkitsevä ($p < 0,0001^{***}$). Kuvassa 7 näkyy sanatyypin vaikutus kohdeäänteeseen ([h, m, p]) normalisoimattomassa ja z-muunnetussa aineistossa.



Kuva 7. Aineiston kohdeäänteet sanatyypeittäin ja äänneittäin. Vasemmalla kuvassa kohdekonsonanttien absoluuttiset kestot, oikealla normalisoidut kestot. Sanatyypit vasemmalta oikealle: epälainasana, epäsanana, nimi, oikea sana. Kunkin sanatyypin kohdalla on kolme laatikkoa, joista vasemmanpuoleinen, vaalein on äänne [h], keskimmäinen [m] ja oikeanpuolimmainen, tummin laatikko on [p].

3.2 Sanan rakenteen vaikutus rajapidennyksen toteutumiseen

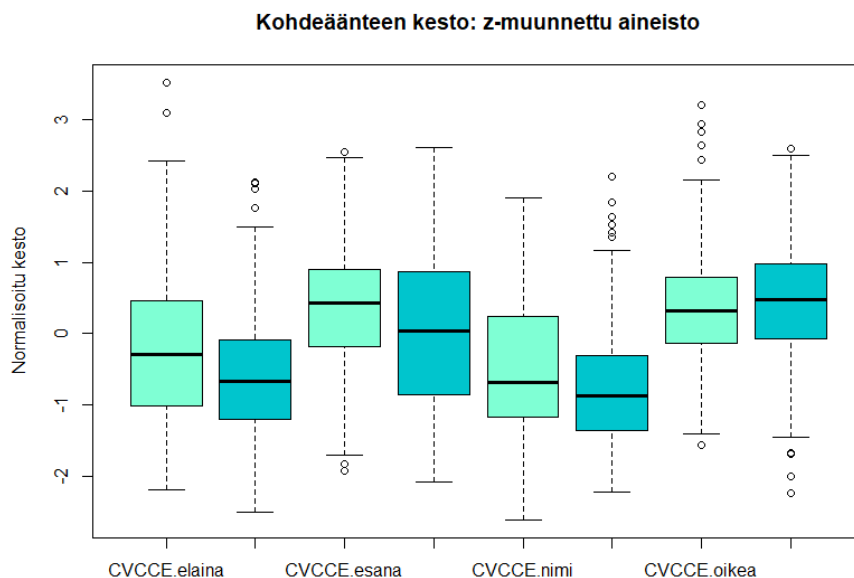
Tutkimuksessa tarkasteltiin myös, vaikuttaako herätesanan rakenne rajapidennyksen toteutumiseen. Sanat olivat rakenteeltaan kahdenmuotoisia: *CVCE tai *CVCCE, joista jälkimmäinen sisälsi geminaatan. Sanan muodon vaikutusta ilmiöön eri sanatyypien yhteydessä tutkittiin varianssianalyysillä ja *post hoc* -testillä Tukey HSD (*honest significant*

difference). Varianssianalyysin epäparametrisena vastineena käytettiin Kruskal-Wallis testia.

3.2.1 Varianssianalyysi: z-muunnettu aineisto

Z-muunnnetulle aineistolle tehdyn varianssianalyysin perusteella herätesanan rakenne oli tilastollisesti merkitsevä tekijä rajapidennyksen toteutumisessa ($p = 0,0028^{**}$). Jotta saataisiin selville, missä sanatyypeissä vaikutusta oli, tehtiin Tukeyn testi. Sanan rakenne osoittautui tilastollisesti merkitseväksi tekijäksi epälainasanojen yhteydessä siten, että *CVCCE-muotoiset, geminaatan sisältävät epälainas sanat aiheuttivat tilastollisesti merkitsevästi pidemmän rajapidennyksen *CVCE-muotoisiin verrattuna ($p = 0,0012^{**}$).

Muissa sanatyypeissä muoto ei vaikuttanut rajapidennyksen toteutumiseen tilastollisesti merkitsevällä tavalla; oikeissa sanoissa p-arvo oli 0,5730 ja nimissä 0,2797. Epäsanoissa p-arvo oli lähimpänä tilastollista merkitsevyyttä ollen 0,0631. Kuvassa 8 näkyy sanan rakenteen vaikutus rajapidennyksen toteutumiseen eri sanatyypeissä normalisoidussa aineistossa.



Kuva 8. Kohdeäänteen z-muunnetyt kestot sanatyyppin ja sanan rakenteen mukaan. Kuvassa kaikki kohdeäänteet ($n = 1872$), äänteittäisiä kestoja ei ole eritelty. Laatikosta vaaleampi, vasemmanpuoleinen on muoto *CVCCE ja oikeanpuoleinen, tummempi on muoto *CVCE. Sanatyytit vasemmalta oikealle: epälainasana, epälasana, nimi, oikea sana.

3.2.2 Kruskal-Wallis test: z-muunnettu aineisto

Vaikka aiempi tilastollinen tarkastelu oli osoittanut tulokset yhteneviksi riippumatta käytetyn menetelmän jakaumaoletuksesta, tehtiin sanan rakenteen vaikutuksen analyysi myös epäparametrisena. Oletetun mukaisesti myös Kruskal-Wallis testin perusteella sanan rakenteella oli tilastollisesti merkitsevä vaikutus kohdeäänteen pituuden toteutumiseen normalisoidussa aineistossa ($p = 0,0076^{**}$).

Normalisoimattomalle aineistolle tehty varianssianalyysi ei ilmentänyt tilastollisesti merkitseviä eroja erirakenteisten herätesanojen välillä merkitsevyysrajan ollessa 0,01. Lähimpänä merkitsevyyden rajaa kuitenkin oli rakenteen vaikutus rajapidennykseen epälainasanoissa, eli tulos mukaili z-muunnetusta aineistosta saatuja tuloksia. Myöskään epäparametrinen testi normalisoimattomassa aineistossa ei antanut tilastollisesti merkitsevää tulosta ($p = 0,0541$).

3.3 Kohdeäänteet

Tutkimuksen kohdeäänteet ovat muuttumattomilta ominaisuuksiltaan hyvin erilaisia. Myös niiden lähtökohtaiset kestot ovat hyvin eripituiset: [p] on soinnittomana klusiilina muita pidempi, [m] ja [h] kestoiltaan melko lähellä toisiaan, mutta [h]:n rajapitenemistä ei oleteta muiden äänteiden lailla. Tämän lisäksi aineistossa on herätesanoja, joiden yhteydessä rajapidennys oletettavasti ilmenee (oikeat sanat), sekä sanoja, jotka sitä luultavimmin eivät aiheuta (nimet). Jokainen äänne esiintyy siis aineistossa todennäköisesti sekä pitkänä että lyhyenä, vaikka lyhyiden ja pitkien lukumääräinen suhde ei oletettavasti olekaan tasainen.

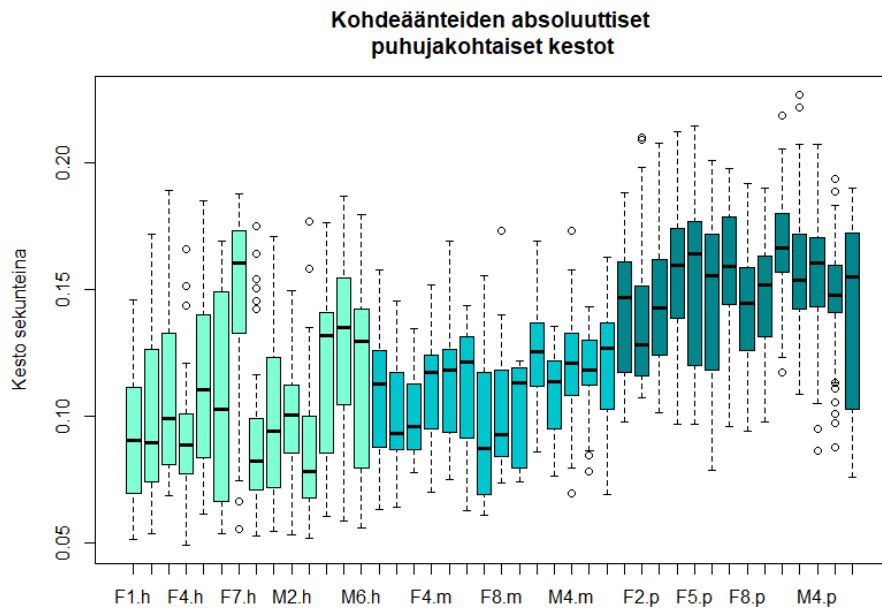
Kohdeäänteitä oli aineistossa alun perin yhtä monta (804 kutakin), mutta lukiessa tapahtuneista virheistä, tauoista ja rytmityksistä johtuvien hylkäämisten vuoksi äänteiden analysoitava määrä ei lopulta ollut aivan tasainen. [h] esiintyi aineistossa lopulta yhteensä 627 kertaa, [m] 587 kertaa ja [p] 658 kertaa.

Taulukosta 5 ilmenevät tämän tutkimuksen kohdeäänteiden keskiarvoiset kestot, keston keskihajonta sekä kestojen mediaanit. [h]:n voi nähdä olevan tässä aineistossa kestoltaan lyhin, vaikka ero sen ja [m]:n välillä on hyvin pieni. Keskihajonta kertoo kuitenkin [h]:n toteutumisessa olevan suurempaa vaihtelua kuin [m]:n, jonka toteutuminen on kohdeäänteistä yhdenmukaisinta. [p] erottuu kahdesta muusta äänneestä selvästi pidemmällä kestollaan, ja sen keskihajonta kertoo suuremmasta vaihtelusta kuin [m]:llä, mutta vähäisemmästä kuin [h]:lla.

Taulukko 5. Aineiston kohdeäänteiden absoluuttisia kestoja.

Äänne	[h]	[m]	[p]
Keston keskiarvo	0,1068 s	0,1111 s	0,1494 s
Keston mediaani	0,1000 s	0,1144 s	0,1525 s
Keskihajonta	0,0341 s	0,0224 s	0,0281 s

Koska kohdeäänteiden kestot eroavat toisistaan jo alkujaan, niiden rajapidennyskestojen vertailu ei olisi kovin mielekästä, semminkin kun myös herätesanan tyyppi ja rakenne vaikuttavat kohdeäänteen toteutumiseen. Normalisoimattomalle aineistolle tehtiin kuitenkin Kruskal-Wallis- ja Wilcoxonin testit, joiden perusteella kaikki kohdeäänteet erosivat toisistaan, kuten oletettua, tilastollisesti erittäin merkitsevästi ($p < 0,0001^{***}$). Taulukkoa 5 noudatellen myös tilastollisen testin mukainen ero [h]:n ja [m]:n välillä oli pienempi kuin [h]:n tai [m]:n ero [p]:hen.



Kuva 9. Koehenkilöiden tuottamat kohdeäänteet normalisoimattomassa aineistossa. F viittaa naispuhujaan, joita oli kahdeksan ja M miespuhujaan, joita oli kuusi. Vasemmanpuoleiset, vaaleimmat laatikot ovat [h]-äänteitä, keskellä ovat [m]:t ja oikeanpuoleiset, tummimmat ovat [p]-äänteitä. Kuvasta on nähtävissä laaja kestopuhtaus koehenkilöiden tuottamien äänteiden välillä. Huomiota kiinnittää puhuja F7, jonka [h]:t ovat huomattavan pitkiä. Hänen puhunnoksaan jouduttiin hylkäämään runsaasti: akustinen analyysi kohdistettiin vain 24 [h]-äänteeseen 60:sta. Näistä lyhyenä toteutuneita oli viisi.

Lähtökohtaisten äännekestoerojen ja koehenkilöiden tuottamien rajapidennyserojen lisäksi kohdeäänteiden kestoissa näkyi myös puhujakohtainen puhenopeus. Kuvassa 9 (yllä) näkyvät tutkimuksen kaikkien koehenkilöiden tuottamat kohdeäänteet absoluuttisissa kestoissaan. Myös kuvasta on mahdollista nähdä, että aineiston lyhimmat äänteet ovat pääosin [h]-äänteitä ja pisimmät [p]-äänteitä. Ero [h]:n ja [m]:n välillä ei kuitenkaan ole suuri.

Koska tutkimuksen pääasiallinen tavoite on selvittää, onko rajapidennyksen toteutuminen pseudosanojen yhteydessä vastaavanlaista kuin oikeiden sanojen – vai vertautuvatko pseudosanat lähinnä nimiin – taulukkoon 6 on koottu kohdeäänteiden keskimääräiset kestot herätesanatyypin mukaan.

Taulukko 6. Kohdeäänteiden keskimääräiset absoluuttiset kestot kunkin herätesanatyypin yhteydessä.

	[h]	[m]	[p]
Epälainasana	0,0990 s	0,1032 s	0,1361 s
Epäsana	0,1141 s	0,1132 s	0,1571 s
Nimi	0,0894 s	0,0968 s	0,1289 s
Oikea sana	0,1159 s	0,1222 s	0,1616 s

Taulukosta 6 ilmenee, että kohdeäänteiden kestot oikeiden sanojen ja epäsanon yhteydessä muistuttavat huomattavasti toisiaan. [h]:n keston ero niiden välillä on jopa häviävän pieni. Myös epälainasana ja nimet mukailevat toisiaan kohdeäänteen toteutumisessa. Kunkin äänteen keskimääräisesti pisin toteutuma on oikean sanan yhteydessä, lyhin nimen yhteydessä.

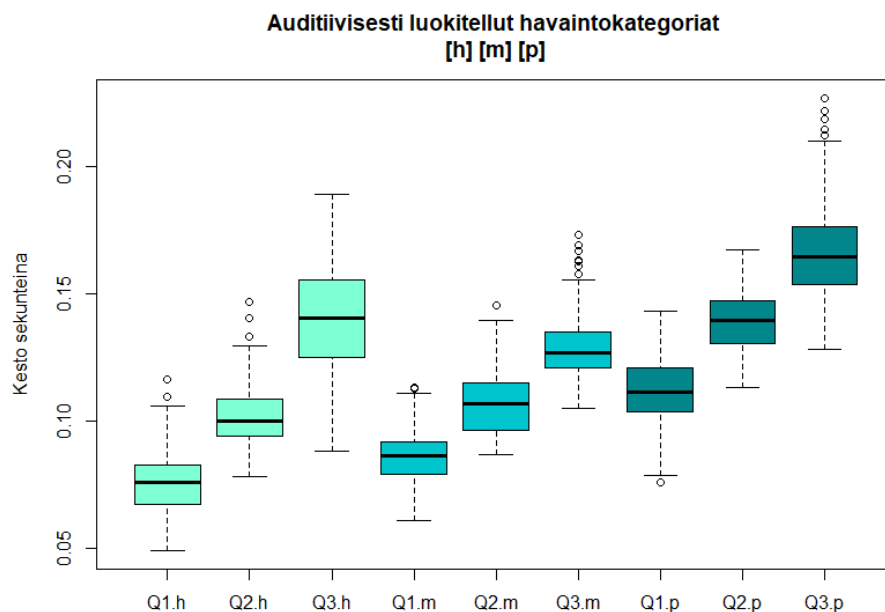
3.4 Auditiivinen luokittelu

Kohdeäänteiden kestoerojen lisäksi haluttiin selvittää, vaikuttaako kohdeäänne siihen, toteutuuko rajapidennys ylipäätään, vai jääkö se toteutumatta. Tähän vaikutti jo aiemmin mainittu [h]:n mahdollinen poikkeuksellisuus rajapidennyksessä. Aineisto oli luokiteltu kuulohavaintojen perusteella kolmeen pidentymisloukkaan sen perusteella, toteutuiko herätesanan jälkeinen kohdeäänne pitkänä (Q3), puolipitkänä (Q2) vai lyhyenä (Q1). Lyhyt (Q1) tarkoitti rajapidennyksen toteutumatta jäämistä. Näiden auditiivisesti luokiteltujen kategorioiden keskimääräiset kohdeäännekohtaiset kestot näkyvät taulukossa 7.

Taulukko 7. Kohdeäänteiden absoluuttisten kestojen keskiarvot auditiivisesti luokitelluissa pituuskategorioissa. Sulkeissa keskihajonta.

Äänne	[h]	[m]	[p]
Q1	0,0757 s (0,0117 s)	0,0863 s (0,0099 s)	0,1121 s (0,0134 s)
Q2	0,1021 s (0,0122 s)	0,1071 s (0,0117 s)	0,1389 s (0,0121 s)
Q3	0,1407 s (0,0213 s)	0,1290 s (0,0122 s)	0,1662 s (0,0167 s)

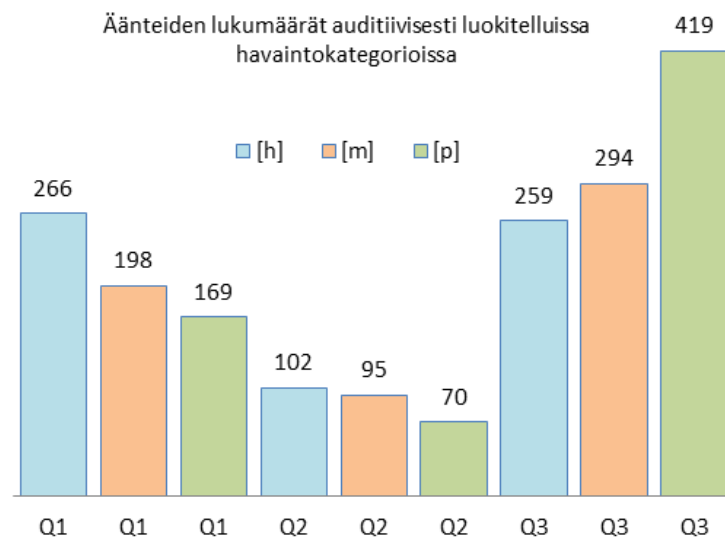
Taulukosta 7 ilmenee, että Q1-kategoriassa, äänten toteutuessa lyhyenä ilman rajapidennystä, [h] on lyhytkestoisin äänneistä, [p] pisin. Puolipitkänä, Q2-luokassa, [h] ja [m] eivät eroa enää mainittavasti toisistaan, ja Q3:ssa, pitkänä reaalistuessaan, [h]:n keskimääräinen kesto ylittää [m]:n keston. Kaikki äänteet pitenevät tasaisesti eri kategorioiden välillä, mutta [m]:n suhteellinen piteneminen on pienempää kuin muiden, joten pisimmillään toteutuessaan se on lyhin luokassaan. Keston yhdenmukaisuudesta ja havaintojen sijoittumisesta ajallisesti lähekkäin kertoo myös [m]:n keskihajonta, joka on vertailun pienin. Kuvassa 10 näkyvät auditiivisesti luokiteltujen kategorioiden kestot normalisoimattomassa aineistossa.



Kuva 10. Kuvassa kohdeäänteiden auditiivisesti luokitellut kestokategoriat. Q1 tarkoittaa lyhyeksi, Q2 puolipitkäksi ja Q3 pitkäksi arvioitua kestoja. Äänteet (n = 1872) vasemmalta oikealle kolmessa pituusluokassaan: [h] (vaalein), [m] (keskimmäinen) ja [p] (tummin).

Jotta selviäisi, onko havaittujen kategorioiden kestojen välillä tilastollisesti merkitsevää eroa, aineistolle tehtiin tilastollinen analyysi. Vaikka aiemmat tulokset olivat osoittaneet tulosten olevan yhteneviä testin jakaumaoletuksesta huolimatta, analyysi tehtiin parametrisena ja epäparametrisena sekä z-muunnetulle aineistolle että absoluuttisille kestoille. Kummallekin aineistolle tehtyjen molempien menetelmien – varianssianalyysin ja Kruskal-Wallis testin – perusteella auditiivisesti luokiteltujen kestokategorioiden välillä oli tilastollisesti erittäin merkitsevä ero ($p < 0,0001^{***}$).

Jotta ilmenisi, eroavatko äänteet toisistaan eri pituuskategorioihin sijoittumisessa, aineistosta piirrettiin pylväskaavio (kuva 11). Kaaviossa ovat mukana kaikki aineiston kohdekonsonantit. Kuvasta voi nähdä, että Q2-luokassa on vähemmän havaintoja kuin muissa luokissa, mutta eri äänteiden edustus siinä on melko tasainen. Äänne [h] esiintyy useammin lyhyenä (Q1) kuin pitkänä (Q3), mikä poikkeaa muista äänneistä, joita on eniten Q3-kategoriassa. Erityisesti [p]:n määrät Q1- ja Q3-luokissa eroavat toisistaan, sillä pitkiä [p]-äänteitä on 2,5-kertaisesti lyhyisiin nähden.

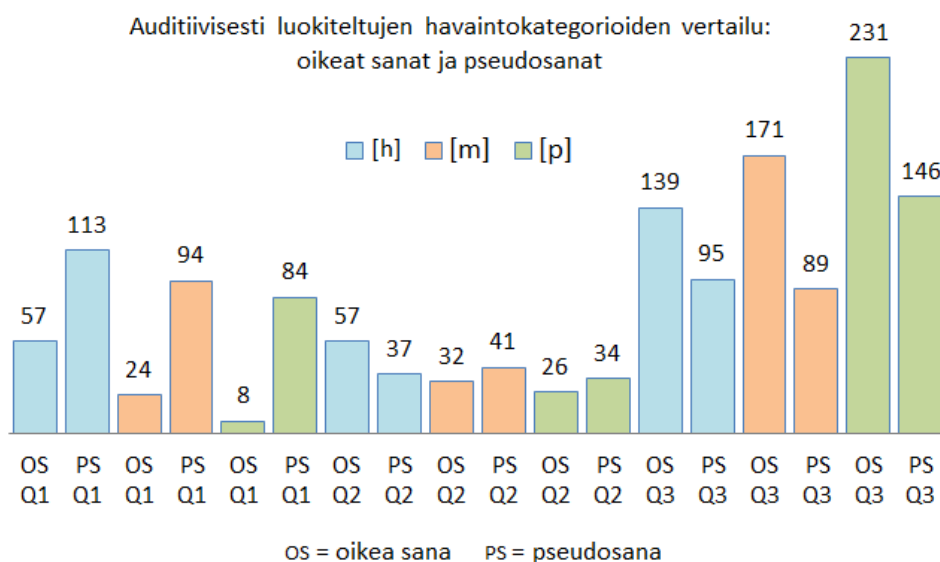


Kuva 11. Pylväskaavio aineiston kaikkien kohdeäänteiden sijoittumisesta kolmeen havaittuun kestokategoriaan. Lukutilanteessa tapahtuneista virheistä johtuvien hylkäämisten vuoksi analysoitavien äänteiden määrä oli hieman epätasainen: [h] $n = 627$, [m] $n = 587$, [p] $n = 658$.

Tutkimuksen pääasiallinen huomionkohde on rajapidennyksen toteutuminen eri sanatyypin yhteydessä, joten sanaryhmiä vertailtiin toisiinsa myös äänteiden kuulonvaraisluokiteltujen kategorioiden osalta. Koelauseet oli laadittu siten, että tietyssä lauseessa esiintyi vain

yhdentyypisiä herätesanoja, mutta kaikki sanat esiintyivät kerran kunkin äänteen kanssa. Äänteiden jakautumisen eri havaintoluokkiin voisi siis olettaa olevan tasainen, jos rajapidennyksen laukaisija olisi pelkkä herätesana. Huomioon on toki otettava myös lausekontekstin ja semanttisen sisällön mahdollinen vaikutus ilmiön toteutumiseen.

Kuvassa 12 oikeiden sanojen ryhmää (OS) on verrattu pseudosanoihin (PS), eli epäsanoihin ja epälainasanoihin yhdessä. Nimet on jätetty vertailun ulkopuolelle. Oikeita sanoja oli tutkimuksessa yhteensä 900 ja pseudosanoja 1008. Hylkäämisten jälkeen oikeita sanoja oli analysoitavana 745 ja pseudosanoja 733 (epäsanoina 378 ja epälainasanoja 355).



Kuva 12. Kohdeäänteiden määrät oikeiden sanojen ja pseudosanojen yhteydessä eri havaintokategorioissa. Kunkin äänteen saman pituusluokan molempien sanaryhmien havainnot ovat vierekkäin: vasemmanpuoleinen on oikea sana (OS), oikeanpuoleinen pseudosana (PS). Analysoitavia äänteitä kaaviossa: [h] n = 498, [m] n = 451, [p] n = 529.

Kuvasta 12 ilmenee, että [h] on äänteistä useimmin lyhyt (Q1) pseudosanojen yhteydessä. Ero [m]:ään ja [p]:hen on kohtalainen, mutta tuntuvampi se on oikeissa sanoissa, joiden yhteydessä [h] toteutuu lyhyenä seitsenkertaisesti [p]:hen nähden. Myös Q2-kategoriassa [h]:sta on äänteistä eniten havaintoja oikeiden sanojen yhteydessä – jopa enemmän kuin pseudosanojen yhteydessä. Kohdeäänne [p]:stä on vähiten havaintoja Q1- ja Q2-luokissa, mikä näkyy runsaana määränä Q3-kategoriassa. [m] on havaintoluokkiin sijoittumisessa määrällisesti [h]:n ja [p]:n välillä, mutta esimerkiksi Q1- ja Q2-luokissa havaintojen suhde oikeiden ja pseudosanojen yhteydessä muistuttaa enemmän [p]:tä kuin [h]:ta.

4 Pohdinta

Rajapidennys on puheen ilmiö, jonka aiheuttavat tietyt vokaaliloppuiset sananmuototyytit, joista tunnusomaisimpia ovat e-loppuiset *vene*-tyypin sanat. Tässä työssä tutkittiin ilmiön foneettista toteutumista sen mahdollisesti laukaisevien herätesanojen avulla, joiden jälkeisen kohdeäänteen kesto segmentoititiin ja mitattiin. Ilmiön produktiivisuutta analysoitiin tutkimusta varten kehitettyjen pseudosanojen avulla. E-loppuisia herätesanatyyppejä oli kaikkiaan neljä: oikeat sanat, epäsanat, epälainasanat ja nimet, ja sanatyytit erosivat toisistaan ensimmäiseltä äänteeltään. Kielen oppimisen ajatellaan yleisesti perustuvan analogiaan, joten rajapidennyksen olisi teoriassa voinut olettaa toteutuvan pseudosanoissa, jotka mukailivat oikeita *vene*-tyypin sanoja. Toisaalta on tiedossa, että *kanne* tyypillisesti laukaisee ilmiön, yhdeltä äänteeltä siitä eroava *Janne* ei. Eri sanatyyppien lisäksi tarkasteltiin sanan rakenteen ja kohdeäänteen laadun vaikutusta ilmiön toteutumiseen.

Rajapidennys osoittautui tämän tutkimuksen perusteella produktiiviseksi ilmiöksi e-loppuisten nominien muototyyteissä, sillä sen aiheuttivat oikeiden *vene*-tyypin sanojen lailla lähes yhtäläisesti myös tutkimusta varten kehitetyt epäsanat. Herätesanatyypeissä oli odotetusti eroja ilmiön laukaisemisessa, sillä pisimmän keskimääräisen keston aiheuttivat oikeat sanat ja lyhimmän nimet. Oikeiden sanojen yhteydessä ilmiö toisin sanoen toteutui useimmin, nimien yhteydessä harvimminkin. Epäsanat muistuttivat ilmiön aiheuttamisessa läheisesti oikeita sanoja ja epälainasanat nimiä. Myös sanan rakenteen havaittiin vaikuttavan epälainasanoissa ilmiön toteutumiseen. Näiden lisäksi kohdeäänteellä todettiin olevan vaikutus sekä rajapidennyksen mitattuun keston että ilmiön toteutumiseen ylipäättään.

E-loppuisten nominien muotoryhmä mainitaan tunnusomaisimpana rajapidennysilmiön ilmentäjänä (Nikkinen-Piraccini, 2011; Matušková, 2016). Syitä tähän voi pohtia. Eräs syistä tunnusomaisuuteen saattaisi olla runsas murrealue- ja puhujakohtainen vaihtelu, sillä Karlssonin ja Lehtosen tutkimustulosten (1977) lisäksi myös Suomi (1996, 75) uskoo jättävänsä rajapidennyksen tuottamatta joidenkin *vene*-tyypin sanojen yhteydessä, toisin kuin vaikkapa yksikön toisen persoonan imperatiivissa. Muotoryhmän rajapidennystendenssi lienee siis tarpeeksi säännönmukainen, jotta kuuliija osaa odottaa sitä ja kiinnittää huomionsa myös siitä poikkeamiseen. Ja koska puhujista ja alueellisista puhetavoista johtuvia poikkeuksia esiintyy runsaasti, ne ovat useimmille tuttuja.

Karlssonin ja Lehtosen (1977) tutkimus on harvoja aiheesta julkaistuja äännekestomittauksia, mutta ilmiön tutkimisen voi nähdä sivuavan kvantiteettitutkimusta. Sananrajainen rajapidennys ei kuitenkaan ole painon lankeamisen vuoksi täysin verrattavissa sanansisäiseen geminaattaan. Tämän lisäksi aineistojen annotoinneissa saattaa olla hienoisia, tutkijakohtaisia eroja, joten tarkkoja vertailuja tutkimusten välillä ei muutenkaan ole tarkoituksenmukaista tehdä.

Aiempien tutkimusten kanssa yhtenevästi soinniton klusiili [p] osoittautui myös tässä tutkimuksessa muita äänneitä pidemmäksi ja altteimmaksi pitenemään, ja [h]:n epäsystemaattisuus rajapidennyksessä ilmeni. Yllättävää oli [m]:n suhteellisen pitenemisen vähyys, mikä eroaa Nahkolan (1987, 14) esittelemistä mittauksista, joissa resonantit pitenevät suhteellisesti obstruentteja enemmän geminoituessaan. Kyseessä saattaisi olla erilaisen annotointitavan lisäksi rajapidennyksen reaalistuminen sanansisäisestä geminaatiosta poikkeavalla tavalla, mihin myös Karlsson ja Lehtonen (1977, 31–33) ovat viitanneet. Myös Laurosela (1922, 17) mainitsee lausefoneettisessa asemassa olevan rajapidennyksen toteutuvan lyhytalkuisena geminaattana, kun lyhyiden vokaalien väliset konsonantit toteutuvat muuten pitkäalkuisina. Eri äänneiden välisistä rajapidennyksen toteutumiseroista ei tämän suhteen kuitenkaan ole mainintoja.

Tulos produktiivisuudesta on hieman vastakkainen Karlssonin ja Lehtosen tutkimustulokseen tekemälle tulkinnalle, jonka mukaan ilmiö olisi poistumassa vähitellen suomesta (Karlsson & Lehtonen, 1977, 15). Myös Karlsson (1982, 350) toteaa, että rajapidennyksestä pyritään eroon useilla murrealueilla, etenkin länsimurteissa, ja Suomi (1996, 75–76) kirjoittaa ilmiön olevan katoamassa ja katoamisen vaikuttavan etenevän etelästä pohjoiseen päin. Edellä mainituista julkaisuista on tosin kulunut jo jonkin aikaa. Ilmiön toteutumista tutkineen Lehtimäen (2012) mukaan, sen sijaan, rajapidennys reaalistuu tietyissä muotoryhmissä jopa enemmän nuoren kuin vanhemman ikäpolven puheessa, mikä viittaisi ilmiön säilymiseen suomessa myös tulevaisuudessa. Myös tämä tutkimus vahvistaa otaksuntaa rajapidennyksen elinvoimaisuudesta.

4.1 Hypoteesien toteutumisen tarkastelu

Tutkimuskysymyksistä tehtyjä hypoteeseja tarkasteltiin sekä parametrisin että epäparametrisin tilastollisin menetelmin, sillä aineisto oli melko suuri, mutta sen normaalijakautuneisuus ei ollut ilmeistä. Aineistolle tehtiin myös normalisointi, jolla häivytettiin koehenkilöiden

puhenopeuseroja, ja aineistoa testattiin sekä normalisoituna että normalisoimattomana. Tilastollisina menetelminä käytettiin lineaarista regressioanalyysiä, varianssianalyysiä sekä varianssianalyysin epäparametrista vastinetta, Kruskal-Wallis testiä Wilcoxonin testin kera. Kaikilla menetelmillä saatiin lähes yhtenevät tulokset, ja tutkimuskysymyksistä tehdyt hypoteesit hyväksyttiin tai hylättiin testien ilmentämien p-arvojen perusteella.

Ensimmäisestä tutkimuskysymyksestä muodostettua nollahypoteesia ”Rajapidennyksen toteutumisessa ei ole eroa sanatyypin välillä” testattiin asettaen vertailukohdaksi herätesanatyypin oikeat sanat. H_0 jäi voimaan epäsanoina, sillä vaikka kohdekonsonantti reaalistui hieman lyhyempikestoisena kuin oikeiden sanojen yhteydessä, ilmiön toteutumisessa oikeiden sanojen ja epäsanoin välillä ei ollut tilastollisesti merkitsevää eroa. Epälainasanat ja nimet puolestaan aiheuttivat tilastollisesti erittäin merkitsevästi lyhyemmän rajapidennyksen oikeisiin sanoihin nähden. Niiden kohdalla H_0 hylättiin hyväksyen vastahypoteesi H_1 : ”Ilmiön toteutuminen on erilaista sanatyypin välillä”.

Myös muita sanatyyppejä vertailtiin toisiinsa. Nimet ja epäsanat erosivat toisistaan tilastollisesti erittäin merkitsevästi, mikä oli odotettua, sillä epäsanat muistuttivat läheisesti oikeita sanoja, joiden ero nimiin oli osoittautunut erittäin merkitseväksi. Nimien ja epälainasanojen välinen ero oli tilastollisesti merkitsevä tai melkein merkitsevä käytetystä menetelmästä ja aineistosta riippuen. Epäsanat ja epälainasanat erosivat toisistaan tilastollisesti erittäin merkitsevästi. Tutkimuksen sanatyypeistä oikeat sanat ja epäsanat muistuttivat toisiaan tilastollisesti eniten rajapidennyksen laukaisemisessa. Toiseksi lähimpänä toisiaan olivat nimet ja epälainasanat.

Tutkimuskysymyksestä 2 tehty nollahypoteesi ”Herätesanan rakenne ei vaikuta rajapidennyksen toteutumiseen” hyväksyttiin muiden sanatyypin osalta, mutta epälainasanoissa se hylättiin hyväksyen vastahypoteesi H_1 : ”Herätesanan rakenne vaikuttaa ilmiön toteutumiseen”. Normalisoiduissa epälainasanoissa geminaatan sisältävä *CVCCE-muoto aiheutti tilastollisesti merkitsevästi pidemmän kohdeäänteen keston kuin *CVCE-muoto.

4.2 Ennakko-oletusten toteutumisen tarkastelu

Tutkimuskysymyksen 1 ennakko-oletuksista toteutuivat jossain määrin kaikki. Rajapidennyksen oletettiin toteutuvan oikeiden sanojen yhteydessä, mikä piti suurimmalta osin paikkansa, mutta toteutuminen ei ollut täysin systemaattista. Osaltaan siihen saattoi vaikuttaa kohdeäänenne, sillä kuten tuloksista ilmeni, [h] toteutui usein lyhyenä, eli ilmiö jäi toteutumatta. Toi-

nen syy saattaisi olla lauseiden semanttinen sisältö. Vaikka lauseita laadittaessa kiinnitettiin runsaasti huomiota siihen, että ne säilyisivät luontevina ja merkityksellisinä myös herätesanan vaihdon jälkeen, saattoi oikea sana silti vaikuttaa jossain kontekstissa yhteyteen sopimattomalta, jolloin sen saattoi mieltää epäuskottavana epäsanana.

Kolmanneksi, myös italian alkukonsonantin vahvistumisen tutkimuksessa huomattiin, että osassa tapauksia, joissa ilmiö olisi voinut toteutua, se jäi toteutumatta (Stevens, Hajek & Absalom, 2002). Syynä ei ollut tauko tai muu ilmiön estävä olosuhde, vaan rajapidennys jätettiin tuottamatta ilman näkyvää syytä. Tutkimuksessa päädyttiinkin otaksuntaan, että pikemmin kuin kategorinen ja teoreettisesti säännöllinen, ilmiö on valinnainen ja epäsäännöllinen. Stevens, Hajek ja Absalom tarkastelivat luonnollista puhetta, joka toki eroaa luku-puheesta, mutta tendenssi saattaisi olla vastaava ja italian päätelmät suomeen rinnastettavia. Myös Karlsson ja Lehtonen (1977, 15) tähdentävät, että rajapidennys ei ole systemaattinen, pakollinen tai ”sataprosenttinen” ilmiö, ja Suomi (1996, 75) kirjoittaa ilmiössä olevan tilannekohtaista vaihtelua.

Ennakko-oletus E1.2 oli, että rajapidennys toteutuisi epäsanojen yhteydessä. Tämän uskottiin tapahtuvan, jos herätesana olisi tarpeeksi uskottava lausekontekstissaan, ja vastavuoroisesti ilmiön oletettiin jäävän toteutumatta, jos epäsanana tuntuisi liian keksityltä. Tämän ennakko-oletuksen voi katsoa pitäneen paikkansa, sillä pääasiassa epäsanana laukaisi rajapidennyksen – lähes vastaavasti kuin oikea sana. Oikeissa sanoissa ilmiö jäi välillä toteutumatta, ja näin oli epäsanojenkin kohdalla. On kuitenkin mahdotonta sanoa, mikä osa toteutumatta jäämisestä on oikeiden sanojen toteutumattomuuden kanssa rinnastettavaa, ja mikä osa selittyy epäsanon epäuskottavuudella tai muulla seikalla. Tämän selvittämiseksi tarvittaisiin lisätutkimusta.

Kolmas ennakko-oletus oli, että epälainasanat eivät todennäköisesti laukaisisi rajapidennystä, sillä niiden oletettiin lainasanojen tavoin lukeutuvan ilmiötä aiheuttamattomaan *nalle*-ryhmään. Suurimmassa osassa tapauksia rajapidennys jäikin toteutumatta epälainasanojen yhteydessä. Toteutuessaan ilmiön aiheuttava epälainasana oli useimmiten *CVCCE-muotoinen, kuten oli otaksuttu.

Ennakko-oletus E1.4 piti paikkansa osittain. Oletus oli, etteivät nimet aiheuta rajapidennystä, ja tämä piti paikkansa oikeiden, tuttujen nimien ja lempinimien kohdalla. Tutkimusta varten keksityt ”epänimet”, kuten *Rumme* ja *Linne*, kuitenkin laukaisivat ilmiön osalla koehenkilöitä. Tämä saattaisi selittyä niiden mahdollisella mieltämisellä sukunimiksi, joiden yhteydessä raja-

pidennys ehkä tuntuisi luontevalta. Kaikilla koehenkilöillä ”epänimet” eivät laukaisseet ilmiötä, mutta sen toteutuessa laukaisija oli tyypillisesti sanarakenteeltaan muotoa *CVCCE.

Ennakko-oletus E2 toiseen tutkimuskysymykseen oli, että sanan rakenne *CVCCE aiheuttaisi rajapidennyksen toteutumisen useammin pseudosanoissa yleisinhimillisen rytmillisyyden tavoittelun vuoksi. Oikeissa sanoissa ja nimissä sanan rakenteella ei uskottu olevan vaikutusta toteutumiseen. Ennakko-oletus piti paikkansa vain osittain, sillä sanan rakenne vaikutti ilmiöön oletetulla tavalla epälainasanoissa, mutta ei epäsanoissa, kuten myös oli arveltu. Kuten oletettua, oikeissa sanoissa rakenteella ei ollut vaikutusta ilmiön toteutumiseen, eikä vaikutus nimissäkään ollut tilastollisesti merkitsevä. ”Epänimien” yhteydessä toteutuessaan rajapidennys kuitenkin toteutui useammin *CVCCE-muodon yhteydessä.

4.3 Koehenkilöt

Tässä tutkimuksessa keskityttiin rajapidennyksen toteutumiseen ilmiönä, joten koehenkilöiden idiolektikohtaisia eroja sen tuotossa ei analysoitu. Aineistoa kuunnellessa ilmeni kuitenkin, että puhujat erosivat tapauskohtaisesti toisistaan. Puhenopeuserojen lisäksi koehenkilöt suhtautuivat eri lailla luettavaan materiaaliin. Osa oli hyvin tietoisia koetilanteesta ja pyrki lukemaan lauseet mahdollisimman huolellisesti, jolloin syntyi helposti taukoja, kun sanat artikuloitiin jopa liioitellun tarkasti. Toiset pyrkivät ohjeistuksesta huolimatta elävöittämään lukemista rytmityksin, mikä myös aiheutti tauon analysoitavaan kohtaan lausetta.

Yhteistä koehenkilöille oli heidän koetilanteen jälkeen kertomansa lukemisen helpottuminen toisella lukukierroksella, minkä voisi mahdollisesti tulkita oppimiseksi tai ainakin tilanteeseen tottumiseksi. Kolmannella kierroksella muutamalla oli puolestaan havaittavissa paikoittaista väsymistä tehtävään. Äänitiedostot on nimetty siten, että ne sisältävät tiedon lukukerrasta, joten potentiaalisia lukukierroseroja olisi mahdollista tarkastella jatkossa myös tilastollisesti.

Osa koehenkilöistä tai heidän vanhempansa olivat kotoisin muualta kuin pääkaupunkiseudulta, mutta kuulohavaintojen perusteella he eivät eronneet toisistaan tai pääkaupunkiseutulaisista puhujista. Poikkeus oli Kotkasta kotoisin oleva informantti, jolla rajapidennys jäi toteutumatta useissa lauseissa, joissa muut koehenkilöt sen pääasiallisesti tuottivat. Kyseisen henkilön spontaanipuheessa vastaavaa toteutumatta jäämistä ei ole kuultavissa, joten kyse lienee osittain fossiloituneesta, Kotkan seudulle ominaisesta artikulaatiotavasta, joka tulee esiin epätavallisessa lukutilanteessa. Puhuja on asunut Helsingissä yli 30 vuotta, mutta lapsuudessa omaksuttujen puheen piirteiden voinee olettaa fossiloituvan puheeseen jokseenkin

samaan tapaan, kuin on osoitettu tapahtuvan vierasta kieltä opittaessa (Seilonen, 2013; Toivola, 2011).

4.4 Muita huomioita tutkimuksesta

Aineistoa annotoitaessa ja auditiivisesti analysoitaessa huomattiin, että lauseen lukemisen rytmityksestä tai artikulaatioselkeyteen pyrkimisestä johtuvia taukoja esiintyi usein huomion kohteena olevalla sananrajalla. Puhujat eivät oletettavasti mieltäneet tekemiään taukoja tauoiksi, eikä tauon käsite ole tutkimuksellisessakaan mielessä täysin yksiselitteinen. Tauon pituudeksi puheessa määritellään yleisimmin 100–400 ms pituiset hiljaiset jaksot, vaikka niitäkin lyhyempiä on raportoitu (Toivola, 2011, 38). Tämän tutkimuksen taukoja ei mitattu, vaan ne merkittiin virheiksi takelteluiden lailla, mutta niiden jatkotutkiminen olisi kiinnostavaa. Hylättyjä lauseita oli kaikkiaan 540, joista oikeita sanoja sisältäviä oli 155 (900:sta) ja pseudosanoja (epäsanat ja epälainat yhdessä) sisältäviä 275 (1008:sta). Vaikka muitakin syitä hylkäämisille oli, selvästi yleisin niistä oli tauko.

Tämä tutkimus ei paneutunut yksilöllisiin eroihin herätesanoissa. Myös tätä olisi kuitenkin kiintoisaa analysoida, sillä joissain tapauksissa epäsanat pitenevät kaikilla lukijoilla ja lähes systemaattisesti kaikissa kolmessa kehyslauseessa, toisten sanojen yhteydessä ilmiö taas jäi säännöllisemmin toteutumatta. Lehtimäen (2012) tutkimuksessa pitkän vokaalin sisältävät epäsanat pitenevät useammin kuin lyhytvokaaliset, ja tämän tutkimuksen lähes systemaattisia pitenejittä olivat esimerkiksi sanat *sunne*, *kinne* ja *humme*. Geminaatan sisältävät epäsanat eivät kuitenkaan osoittautuneet tilastollisesti merkitsevästi useammin ilmiön aiheuttajiksi yksikonsonanttisiin nähden. Esimerkkinä rajapidennyksen merkityksiä erottavasta ominaisuudesta (Karlsson & Lehtonen, 1977, 26) ilmeni koetilanteessa luettu herätesana *murre*, joka ilman rajapidennystä lausuttuna toi mieleen *murteen* sijaan *Murren*.

Kohdeäänteen vaikutusta ilmiön toteutumiseen tai toteutumatta jäämiseen tarkasteltiin auditiivisesti luokiteltujen kategorioiden avulla. Eriteltyjä sanatyypikohtaisia vertailuja ei kuitenkaan tehty, vaan oikeita sanoja verrattiin epäsanoihin ja epälainasanoihin yhdessä. Tätä olisi myös kiinnostavaa tutkia tarkemmin, sillä tässä tutkimuksessa [p] pitenei muita äänteitä useammin, mikä mukailee Liekon (1994) havaintoja lapsen rajapidennyksen tuottamisesta. Lauseen semanttista sisältöä ei voi kuitenkaan sulkea pois ilmiöön mahdollisesti vaikuttavien muuttujien listalta, sillä on mahdollista, että verbi *patistaa* laukaisee ilmiön tyypillisemmin kuin *huokaisi* tai *muisteli*.

Suomen [h] mainitaan artikulaatiopaikaltaan vaihtelevaksi, ja sillä on monta allofonia äänneympäristöstä riippuen. Tässä tutkimuksessa kohdeäänne oli aina vokaalien välinen, joten [h] toteutui pääasiassa soinnillisena. Kaksi puhujaa tuotti äänten kuitenkin usein soinnittomana, mikä aiheutti auditiivisesti arvioituna erityisen pitkän keston. Ehdotusta [h]:n laadullisen eron arvioimisesta pituuden sijaan (Penttilä, 1963, viitattu lähteessä Itkonen, 1964a, 57) voisi kenties muokata tutkimuskysymykseksi siitä, onko [h]:n laadulla ja kestolla korrelaatiota, vai onko ero ainoastaan kuulohavainnossa. Kuulokuvaan pituudesta vaikuttavat toki myös intensiteetti ja painotus, mutta aihe on mielenkiintoinen.

Kohdeäänteen lisäksi aineistoa annotoitaessa myös herätesanan vokaali rajattiin kohdeäännekerrokselle. Korrelaatiotestin perusteella konsonantin ja vokaalin kestolla ei ollut yhteyttä normalisoidussa aineistossa, mutta aihetta ei tutkittu tarkemmin, esimerkiksi eri sanatyypin yhteydessä. Myös tätä olisi mielenkiintoista tarkastella lisää.

4.5 Tutkimuksen arviointia

Puheen tutkimus olisi autenttisinta tehdä spontaanipuheesta. Tässä työssä tarkasteltiin rajapidennyksen toteutumista erilaisten herätesanojen avulla, ja koska herätesanoina käytettiin myös pseudosanoja, luonnollisen puheen tarkastelu ei olisi ollut mahdollista. Tämän tutkimuksen lukupuhunnasta analysoitujen tulosten voinee kuitenkin olettaa vertautuvan ainakin suuntaa-antavasti myös spontaanipuheeseen.

Tutkimuksessa käytetty koehenkilöiden valintamenettely ei ole suositeltavin, sillä kaikki informantit olivat tutkimuksen tekijälle tuttuja, joten heissä voisi olettaa olevan jotakin yhteistä. Tällöin he eivät täysin edusta perusjoukkoa. (mm. Gries, 2013, 160.) Valintatapaa puoltaa kuitenkin se, että tässä tutkimuksessa olennaista oli ilmiön toteuttaminen tietyssä muototyypissä spontaanipuheessa, ja satunnaisotannalla valituista henkilöistä osa olisi mahdollisesti jouduttu toteuttamattomuuden vuoksi rajaamaan ulos tutkimuksesta. Koehenkilöistä kahdeksan oli naisia ja kuusi miehiä, joten tässä mielessä puhujat ovat pienestä määrästä huolimatta suhteellisen yhtäläisesti perusjoukkoon yleistettäviä.

Tutkimuksen toteutus pyrittiin suunnittelemaan huolellisesti, ottamaan huomioon mahdollisimman monta muuttujaa ja neutraloimaan ei-tutkittavien muuttujien vaikutusta ilmiöön. Tutkimuskohteen paljastumisen estämiseksi koelauseista tehtiin aihepiiriltään vaihtelevia, mahdollisimman uskottavia ja monisanaisia, mutta helposti yhdellä hengityksellä luettavia. Osa puhujista kuitenkin rytmitti lukuaan tavalla, joka aiheutti tauon ennen kohdeäännettä,

mikä johti lauseen hylkäämiseen. Lauseenvastikkeella alkavissa lauseissa tauko toteutui luonnostaan lauseenvastikkeen jälkeen, jolloin puhe oli yhtäjaksoista analysoitavalla sananrajalla. Tätä olisi mahdollista hyödyntää myös vastaisuudessa rytmittämään lukua. Koska tutkimuksen tarkoitus ei paljastunut yhdellekään koehenkilöistä, on mahdollista jälkeinpäin sanoa myös, että lyhyemmillä lauseilla olisi kenties saatu vältettyä osa tauoista.

Tarkastelukohteena oli kolmen eri kohdeäänteen toteutuminen ja mahdollinen vaikutus rajapidennyksen toteutumiseen. [p], [m] ja [h] ovat lähtökohtaisesti hyvin eripituisia, mikä oli tiedossa tutkimusta suunniteltaessa. Aineistoa annotoitaessa huomattiin, että vaikka klusiili ja spirantti ovat selkeitä rajata puhesignaalista, nasaalikonsonantti ei sitä aina ole. Nasaalisuus piirteenä leviää usein sitä seuraavaan vokaaliin, minkä vuoksi [m]:n rajoja oli haastavampaa saada yhteneväisiksi (ks. myös Campbell, 2000, 303). Jos työssä oltaisi keskitytty vain siihen, laukaiseeko tietty sana ilmiön, olisi ollut perusteltua käyttää enemmän toistensa kaltaisia, tai jopa samaa kohdeäännettä. Tällä olisi tosin saattanut olla monotonistava vaikutus lukemiseen, tai tutkimuksen tarkoitus olisi saattanut valjeta jollekin informanteista. Äännerajojen yhdenmukaisuutta kontrolloitiin tarkastamalla koko aineisto annotoinnin valmistumisen jälkeen ja tekemällä tarvittavat tarkennukset.

Konsonantin foneettiseen kestoon sanassa vaikuttaa sitä edeltävän ja sen jälkeisen vokaalin kesto eli onko vokaali lyhyt vai pitkä (Nahkola, 1987, 14). Vaikka lähteistä ei tiettävästi ilmene vastaavaa vokaalin laadun vaikutusta konsonanttiin – konsonantin laadun vaikutus vokaaliin kylläkin (Lehiste, 1970, 20) – saattaa ääntöväylän muuntelu olla erikestoista liikuttaessa [e]:stä [i]:hin tai [ɑ]:han. Jatkotutkimuksessa kohdesanan voisi vakioda siten, että myös kohdesegmenttiä seuraava vokaali olisi aina sama, herätesanahan on siten vakioitu, että se päättyy aina [e]:hen.

Tässä tutkimuksessa käytettiin tutkijan auditiivisia havaintoja yhtenä muuttujana ilmiötä analysoidessa. Tutkimuksen tekijä on äidinkielen suomenpuhuja ja tottunut havainnoimaan kvantiteettia. Havainnot ovat kuitenkin aina jossain määrin subjektiivisia, joten tutkimuksen sisäistä reliabiliteettia pyrittiin vahvistamaan tekemällä auditiivinen luokittelu useampaan kertaan. Karlsson ja Lehtonen (1977) käyttivät aineistonsa (jossa kohdesanat olivat sekä vokaali- että konsonanttialkuisia) auditiivisessa luokittelussa neljää havaintoluokkaa, ja Nahkola (1987) on kyennyt erottamaan konsonantin kestolle viisi astetta: lyhyt, puolipitkä ja vajaapitkä konsonantti sekä lyhytaltuinen ja täysi geminaatta. Hänen mukaansa viisijako ”ei ole ylivoimainen korvan diskriminaatiokyvylle eikä toisaalta liian karkea foneettisten

väliasteiden luokitteluun” (Nahkola, 1987, 10). Myös Laurosela (1922, 17) erottaa konsonantin kestolle viisi astetta alilyhyestä geminaataan. Tässä tutkimuksessa pitäydittiin kuitenkin kolmijaossa.

4.6 Mietteitä jatkotutkimuksesta

Tällä tutkimuksella osoitettiin rajapidennyksen toteutumisessa olevan huomattavaa vaihtelua eri herätesanojen yhteydessä. Luvuissa 4.3–4.5 mainittujen jatkotutkimusajastusten lisäksi aihetta olisi mahdollista syventää jatkotutkimuksella e-loppuisista epäsanoista käyttäen niissä erilaisia sanarakenteita tai muotoilemalla osa mukailemaan *nalle*-sanoja. Tämän tutkimuksen havaintoja koelauseiden muodostamisesta ja ohjeistuksesta lukutilanteessa olisi mahdollista hyödyntää myös jatkossa.

Erittäin mielenkiintoinen tutkimuskohde olisi rajapidennyksen havaitseminen. Ilmiön havaitsemista on kartoitettu kyselytutkimuksilla, mutta kuuntelukokeita ei liene toteutettu montaa. Karlsson ja Lehtonen (1977, 26–29) tarkastelivat myös ilmiön havaitsemista. Heidän tutkimuksensa turkulaiset koehenkilöt eivät pääasiassa toteuttaneet rajapidennystä e-loppuisissa nomineissa, ja heidän kyseisen muotoryhmän havaitsemisessaan olikin huomattavissa vieraan kielen oppimiseen rinnastettavaa adaptoitumista kokeen edetessä. Havaitsemista olisikin mielenkiintoista tutkia sekä ilmiön toteuttavilla että toteuttamatta jättävillä puhujilla.

Äidinkielisten puhujien lisäksi rajapidennyksen havaitsemista ja tuottamista olisi kiinnostavaa tutkia suomea opiskelevilla henkilöillä, sillä S2-opetuksen tarve lisääntyy jatkuvasti.

Kieliopin lisäksi opiskeltavan kielen suullinen taito on tärkeässä osassa, ja kvantiteetin on raportoitu olevan eräs haasteista. Vaikka rajapidennyksen toteutumattomuus suomen opiskelijoilla on hyväksyttävää (Aho, Huhtaniemi & Nikonen, 2016, 68), myös sen tärkeys S2-opetuksessa muun prosodian ohella on mainittu (Lauranto & Vehkanen, 2014, 79).

Rajapidennyksen voikin nähdä osana suomen kvantiteetti-ilmiötä, jonka taitaminen sekä parantaa ymmärrettävyyttä että lisää idiomaattisuutta puheeseen.

5 Yhteenveto

Tässä foneettisessa tutkimuksessa tarkasteltiin rajapidennyksen toteutumista e-loppuisten herätesanojen avulla. Tutkimuksen pääasiallinen tavoite oli selvittää, onko ilmiö produktiivinen, ja tätä analysoitiin vertaamalla e-loppuisten *vene*-tyypin sanojen ja tutkimusta varten kehitettyjen, *vene*-sanoja mukailevien pseudosanojen rajapidennyksen aiheuttamista suomea äidinkielenään puhuvien koehenkilöiden lukupuhunnassa. Herätesanatyyppejä oli kaikkiaan neljä: oikeat sanat, epäsanat, epälainasanat ja nimet. Sanatyypit erosivat toisistaan vain ensimmäiseltä äänteeltään, joskin lainasanamainen miellelyhtymä tuotettiin vierasperäisen konsonantin lisäksi sananalkuisella konsonanttiklusterilla. Eri sanatyyppejen lisäksi tarkasteltiin sanan rakenteen ja kohdeäänteen laadun vaikutusta ilmiön toteutumiseen.

Herätesanatyypeissä ilmeni odotetusti eroja ilmiön aiheuttamisessa. Oikeiden sanojen yhteydessä rajapidennys toteutui useimmin, nimien yhteydessä harvimminkin. Epäsanat osoittautuivat ilmiön laukaisemisessa läheisesti oikeiden sanojen kaltaisiksi, epälainasanat muistuttivat puolestaan nimiä. Sanan rakenteen havaittiin vaikuttavan ilmiön toteutumiseen puhenopeusnormalisoitujen epälainasanojen yhteydessä siten, että geminaatan sisältävä epälainasana aiheutti rajapidennyksen useammin kuin yksikonsonanttinen. Myös kohdeäänteellä todettiin olevan vaikutus – sekä rajapidennyksen mitattuun keston että ilmiön toteutumiseen ylipäättään.

Rajapidennys osoittautui tämän tutkimuksen perusteella produktiiviseksi ilmiöksi, sillä sen aiheuttivat oikeiden *vene*-tyypin sanojen lailla lähes yhtäläisesti myös tutkimusta varten kehitetyt kotoperäisistä äänneistä muodostuvat epäsanat. Ilmiön produktiivisuudessa ja yhdenmukaisuudessa saattaa toki olla eroja eri muototyypeissä, mutta sen tunnusomaisin muotoryhmä, e-loppuiset nominit, vaikuttaa tämän tutkimuksen perusteella elinvoimaiselta. Rajapidennyksen paikoittainen toteutumattomuus voidaan siis tulkita pikemmin tapaus-, murre- tai puhujakohtaiseksi vaihteluksi kuin osaksi käynnissä olevaa kielenmuutosta.

Lähteet

- Absalom, M., Hajek, J. & Stevens, M. (2004). Il fenomeno del raddoppiamento sintattico nella realtà linguistica italiana. *Proceedings of Il Parlato Italiano*. Napoli: M. D'Auria Editore.
- Aho, E., Huhtaniemi, A. & Nikonen, M. (2016). *Fonetiikkaa suomen kielen oppijoille*. Helsinki: Finn Lectura.
- Aho, E., Toivola, M., Karlsson, F. & Lennes, M. (2016). Aikuisten maahanmuuttajien suomen ääntämisestä. *Puhe ja kieli* 36(2), 77–96.
- Becker von, R. (1824). *Finsk grammatik*. Turku: Bibel-sällskapets tryckeri.
- Boersma, P. & Weenink, D. (2019). Praat: doing phonetics by computer [tietokoneohjelma, versio 6.1.07]. Haettu 28.11.2019 osoitteesta <http://www.praat.org/>
- Campbell, N. (2000). Timing in speech: A multi-level process. Teoksessa M. Horne (toim.), *Prosody: Theory and experiment* (s. 281–334). (Kirjasarjassa Text, speech and language technology). Dordrecht: Springer.
- Dryer, M. S. & Haspelmath, M. (toim.) (2013). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
Haettu 19.8.2020 osoitteesta <http://wals.info/>
- Gordon, M. K. (2016). *Phonological typology*. Oxford University Press.
- Gries, S. T. (2013). *Statistics for linguistics with R. A practical introduction. (2nd revised edition)*. Berliini: De Gruyter Mouton.
- Hakulinen, L. (1979). *Suomen kielen rakenne ja kehitys. Neljäs, korjattu ja lisätty painos*. Helsinki: Otava.
- Häkkinen, K. (1994). *Agricolasta nykykieleen: suomen kirjakielen historia*. Juva: WSOY.
- Ikola, O. (1953). Jäännöslopuke. *Virittäjä*, 57, 431.
- Ikola, O. (1992). *Nykysuomen käsikirja*. Jyväskylä: Weilin+Göös.
- Ikola, O. (2001). *Nykysuomen opas*. (Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 65). Turun yliopisto.
- Itkonen, T. (1964a). *Proto-Finnic final consonants: their history in the Finnic languages with particular reference to the Finnish dialects. I:1. Introduction: the history of -k in Finnish*. Helsinki: Suomalaisen Kirjallisuuden Kirjapaino Oy.
- Itkonen, T. (1964b). Sananrajaisten äänneilmiöiden synkroniaa ja diakroniaa. *Virittäjä*, 68(3), 225–232.
- Itkonen, T. (1969). Loppukahdennus, ei "alkukahdennus". *Virittäjä*, 73, 212–214.

- Itkonen, T. (1987). Erään vokaalivyyhdän selvittelyä. *Virittäjä*, 91(2), 164–208. Haettu 4.5.2020 osoitteesta <https://journal.fi/virittaja/article/view/38104>
- ISK = A. Hakulinen, M. Vilkuna, R. Korhonen, V. Koivisto, T.R. Heinonen & I. Alho. (2004). *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Karlsson, F. (1969). Suomen yleiskielen segmentaalifoneemien paradigma. *Virittäjä*, 73(4), 351–362. Haettu 15.9.2020 osoitteesta <https://journal.fi/virittaja/article/view/35661>
- Karlsson, F. (1973). *Mitä suomen ”loppukahdennus” on?* Turun yliopiston fonetiikan laitoksen julkaisuja, Sarja B, Moniste 2. Turun yliopisto.
- Karlsson, F. (1982). *Suomen kielen äänne- ja muotorakenne*. Juva: WSOY.
- Karlsson, F. & Lehtonen, J. (1977). *Alkukahdennus. Näkökohtia eräistä suomen kielen sandhi-ilmiöistä*. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 2. Turun yliopisto.
- Kielitoimiston ohjepankki, verkkolähde, haettu 29.9.2020 osoitteesta www.kielitoimistonohjepankki.fi/haku/nimi/ohje/185
- Kisler, T., Reichel U. D. & Schiel, F. (2017): Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326–347.
- Koivisto, V. (2013). *Suomen sanojen rakenne*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Laaksonen, K. & Lieko, A. (1988). *Suomen kielen äänne- ja muoto-oppi*. Helsinki: Finn Lectura.
- Lauranto, Y. & Vehkanen, M. (2014). *Päättymätön projekti: puhetta eri S2-foorumilla*. (Julkaisusarjassa Kakkoskieli 7). Helsingin yliopisto.
- Laurosela, J. (1922). *Foneettinen tutkimus Etelä-Pohjanmaan murteesta*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Lehiste, I. (1970). *Suprasegmentals*. Massachusetts: The M.I.T. Press.
- Lehtimäki, M. (2012). *Jäännösloppukkeen toteutuminen nykysuomessa kyselytestimenetelmän avulla tarkasteltuna*. Pro Gradu -tutkielma. Tampereen yliopisto.
- Lennes, M., Aho, E., Toivola, M. & Wahlberg, L. (2006). On the use of the glottal stop in Finnish conversational speech. *Fonetiikan päivät 2006*, 93–102. Helsingin yliopiston puhetieteiden laitoksen julkaisuja 53. Helsingin yliopisto.
- Lieko, A. (1994). Loppukahdennuksen oppiminen. Tapaustudkimus lapsen kielestä. *Virittäjä*, 98(3), 406–429.
- Matušková, L. (2016). *Rajageminaatio – yksi suomen kielen sandhi-ilmiöistä tavallisten kielenkäyttäjien näkökulmasta*. Pro Gradu -tutkielma. Brno: Masaryk university.
- Mielikäinen, A. (2005). Matkimuksista määritelmiin. Miten murteista puhutaan. *Sananjalka*, 47(1), 98–118. Haettu 3.3.2020 osoitteesta <https://doi.org/10.30673/sja.86658>

- Mielikäinen, A. & Palander, M. (2014). *Miten suomalaiset puhuvat murteista? Kansanlingvistinen tutkimus metakielestä*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Nahkola, K. (1987). *Yleisgeminaatio: ääntenmuutoksen synty ja vaiheet kielisysteemissä erityisesti Tampereen seudun hämäläismurteiden kannalta*. Väitöskirja. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Nikkinen-Piraccini, J. (2011). *Keskustelua hernekeitosta: verkkokeskustelijoiden käsityksiä rajageminaatiosta*. Pro Gradu -tutkielma. Helsingin yliopisto.
- Nummenmaa, L. (2004). *Käyttäytymistieteiden tilastolliset menetelmät*. Helsinki: Tammi.
- Penttilä, A. (1969). Vielä alkukahdennuksesta. *Virittäjä*, 73(4), 396–399.
- Raimo, I. & Ojala, S. (2009). Akustiikkaa ja artikulaatioita eli miltä puhe näyttää. Teoksessa O. Aaltonen, R. Aulanko, A. Iivonen, A. Klippi & M. Vainio (toim.), *Puhuva ihminen. Puhetieteiden perusteet* (s. 174–182). Helsinki: Otava.
- Rapola, M. (1966). *Suomen kielen äännehistorian luennot*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- R Core Team (2020). R: A language and environment for statistical computing. [tietokoneohjelma]. R Foundation for Statistical Computing, Vienna, Austria. Haettu 7.10.2020 osoitteesta <https://www.R-project.org/>.
- Renvall, G. (1840). *Finsk språklära: enligt den rena Vest-Finska, i bokspråk vanliga dialecten*. Turku: Christian Ludvig Hjelt.
- RStudio Team (2020). RStudio: Integrated Development Environment for R [tietokoneohjelma]. RStudio, PBC, Boston, MA. Haettu 7.10.2020 osoitteesta <http://www.rstudio.com/>.
- Räisänen, A. (1975). Havaintoja lastenkielestä. *Virittäjä*, 79(3), 251–266. Haettu 10.11.2019 osoitteesta <https://journal.fi/virittaja/article/view/36611>
- Schiell, F. (1999). Automatic phonetic transcription of non-prompted speech. *Proceedings of the ICPhS*, 607–610.
- Seilonen, M. (2013). *Epäsuora henkilöön viittaaminen oppijansuomessa*. Väitöskirja. Jyväskylän yliopisto.
- Setälä, E. N. (1899). *Yhteissuomalainen äännehistoria*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Stevens, M., Hajek, J. & Absalom, M. (2002). Raddoppiamneto sintattico and glottalization phenomena in Italian: a first phonetic excursus. *Proceedings of 9th Australian International Conference on Speech Science and Technology*, 154–159.
- Suomi, K. (1996). *Fonologian perusteita*. Oulun yliopisto.
- Suomi, K. (2009). Minimaalisista ja muista lyhyistä puhunnoksista. *Virittäjä*, 113(1), 58–76. Haettu 12.05.2019 osoitteesta <https://journal.fi/virittaja/article/view/4165>

Suomi, K., Toivonen, J. & Ylitalo, R. (2006). *Fonetiikan ja suomen äänneopin perusteet*. Helsinki: Gaudeamus.

Tieteen termipankki, verkkolähde, haettu 14.3.2019 osoitteesta
<https://tieteentermipankki.fi/wiki/Kielitiede:sandhi>

Tieteen termipankki, verkkolähde, haettu 30.12.2019 osoitteesta
<https://tieteentermipankki.fi/wiki/Nimitys:epäsana>

Toivola, M. (2011). *Vieraan aksentin arviointi ja mittaaminen suomessa*. Väitöskirja. Helsingin yliopiston käyttäytymistieteellisen laitoksen julkaisuja. Helsingin yliopisto.

Treccani, verkkolähde, haettu 25.1.2020 osoitteesta
www.treccani.it/enciclopedia/raddoppiamento-sintattico_%28La-grammatica-italiana%29/

Vainio, M. (2001). *Artificial neural network based prosody models for Finnish text-to-speech synthesis*. Väitöskirja. Helsingin yliopiston fonetiikan laitoksen julkaisuja. Helsingin yliopisto.

Vainio, M., Järvikivi, J., Aalto, D. & Suni, A. (2010). Phonetic tone signals phonological quantity and word structure. *The Journal of the Acoustical Society of America*, 128(3), 1313–1321.

Verkkokielioppi, verkkolähde, haettu 28.3.2019–30.12.2019 osoitteesta
<https://fl.finnlectura.fi/verkkosuomi/Fonologia/sivu191.htm>

WALS online, verkkolähde, ks. Dryer & Haspelmath (toim.) (2013).

Wiik, K. (1998). *Fonetiikan perusteet*. Helsinki: WSOY.